Columbia University

Contributions to Education

Teachers College Series

THE OVERLAPPING OF ATTAINMENTS IN CERTAIN SIXTH, SEVENTH, AND EIGHTH GRADES

BY
PAUL J. KRUSE, PH.D.

TEACHERS COLLEGE, COLUMBIA UNIVERSITY
CONTRIBUTIONS TO EDUCATION, No. 92

PUBLISHED BY
Teachers College, Columbia University
NEW YORK CITY
1918

THE HERSHEY PRESS, HERSHEY, PA.

ACKNOWLEDGMENTS

It is only through the co-operation of those interested in the promotion of the science of education that studies of this nature are possible. I am glad to acknowledge my indebtedness and record my thanks to those teachers in the Seattle Public Schools whose intelligent co-operation made the giving of the tests far less difficult than it would otherwise have been. To the genuine professional interest and ever courteous help of Supt. F. E. Willard and Principals A. S. Gist, W. H. Ellert, R. W. Moore and Charles W. Metsker I owe much. I am unable to express even a small part of my great obligation and deep gratitude to Professor E. L. Thorndike for his helpful advice and encouragement throughout this investigation, and for his generous giving of valuable time and help on other occasions. Without the encouraging assistance given by my wife through the months of testing, scoring, and calculating, the completion of the study at this time would have been impossible.

P. J. K.

CONTENTS

Chapter		Page
I	Introduction	5
	1 Evidences of Overlapping	
	2 The Problems	
II	The Data	8
	1 The Subjects	
	2 Tests Used	
	3 Administration of the Tests	
	4 Scoring the Tests	
III	Overlapping by Single Tests	30
	1 The Measure of Overlapping	
	2 Elimination of Some Tests	
	3 Comparison of Different Tests as to Amount of Overlapping	
	4 Surfaces of Overlapping	
IV	Overlapping in the Case of Composites	42
	1 Results	
	2 Surfaces of Overlapping	
	3 Statistical Treatment	
V	Reliability of the Tests Used as Determined by the Correlation of Groups of Similar Tests	57
VI	Comparison of Results in Overlapping With Those of Other Investigators	59
VII	Conclusions	67
VIII	Bibliography	69
IX	Appendix	71

THE OVERLAPPING OF ATTAINMENTS IN CERTAIN SIXTH, SEVENTH, AND EIGHTH GRADES

I

INTRODUCTION

1. Evidences of Overlapping

One of the most important findings resulting from the measurement of the achievements of pupils by means of standard tests is the great variability in a given grade. This brings in question our whole system of grading, which proceeds upon the assumption that we have in a given grade pupils differing markedly in ability to do the work of the school from those in the grades above and below. There is a tacit understanding that the presence of a pupil in the seventh grade, for example, means that his ability is greater than that representative of the sixth grade and less than that of the eighth grade. The results of much of the testing taken at face value seem to challenge very seriously the validity of these assumptions.

Some of the findings of the Butte Survey Commission are indicated in the following extracts. (14)

"A very wide range of ability in each grade is revealed (in composition) . . . One eighth-grade pupil wrote a composition rated 0 while two pupils wrote papers rated 7 on the scale of 10. The eighth-grade group range rather evenly over all the steps of the scale from 1 to 6 . . . We see that some fourth-grade pupils surpass the median ability of the eighth-grade pupils, while many eighth-grade pupils fall below the median ability of the fourth-grade pupils." (p. 74) "As in spelling and composition, a very wide range of abilities in penmanship is found in each grade . . . . Some children in the second grade surpass the ability of the median eighth-grade pupil, while some in the eighth grade fall below the median second-grade pupil." (p. 81)

Dr. Daniel Starch in his "Educational Measurements" (16) makes repeated reference to this overlapping of grades. "There are pupils in the fourth grade whose attainment in reading is higher than that of the average eighth-grade pupil. Likewise, there are pupils in the fourth grade whose attainment in reading is inferior to that of the average first-grade pupil." (p. 40) "In speed and comprehension combined, 31.8 per cent of the pupils of any grade reached or exceeded the median of the next grade above, 20.1 per cent reached or exceeded the median of the second grade above, 13.2 per cent reached or exceeded the median of the third grade above, and 3.3 per cent reached or exceeded the median of the fourth grade above." (p. 42) "We find that in quality of writing on the average 37.1 per cent of the pupils of any given grade reach or exceed the median of the next grade above it, 24.0 per cent reach or exceed the median of the second grade above it, 14.6 per cent reach or exceed the median of the third grade above it, and 7.7 per cent reach or exceed the median of the fourth grade above it. Statements of the same sort apply to the speed of writing." (pp. 86, 88) "The spelling tests reveal identically the same facts concerning the distribution of the pupils and the overlapping of the grades as were found in case of the reading and writing tests . . . . . The overlapping among the various grades is enormous. There are two pupils in the second grade who can spell as well as two pupils in the eighth grade . . . On the average 20.8 per cent of the pupils of any grade reach or exceed the median of the next grade above it, and 6.2 per cent reach or exceed the median of the second grade above it." (p. 98)

2. The Problems

These statements suggest many questions the answers to which should be highly significant for educational theory and school practice.

a. Will further investigation under carefully controlled conditions support these findings?

b. Have we in the results from single tests in different traits a valid measure of the overlapping of general ability to do the work of a given grade?

c. Is our grading system, in which we have had so much faith, really as bad as it appears to be as a means of selecting on the basis of achievement?

d. There is a tacit feeling that the passing from one grade to the next —with all the tribulation which it entails on the part of pupil and teacher—means a very definite step in accomplishment. Is this feeling well founded?

e. What constitutes a valid measure of overlapping?

It was in the hope of presenting some facts that would have a direct bearing on these and similar questions that this study was undertaken.

Primary Problem. To determine from data, adequate and reliable, the amount and nature of the overlapping in certain sixth, seventh, and eighth grades.

Related Problem. To get a measure of the reliability of certain tests as measures of attainment by school pupils in these grades.

II

THE DATA

It was recognized at the beginning that whatever value might come from such a study as the one contemplated would depend very largely upon (*a*) the number of pupils tested, (*b*) the number and worth of the tests used, (*c*) the uniformity of procedure in giving the tests, (*d*) the opportunity given the pupils to show their real capacity in each test, (*e*) the consistency in the scoring of the papers.

1. The Subjects

The pupils tested were all the sixth-, seventh-, and eighth-grade pupils in four public schools in Seattle, Washington. About 860 pupils in all were tested, this being the number represented in the final composite: 254 in the sixth grade, 324 in the seventh, and 282 in the eighth. The discrepancy between these figures and the totals for certain tests is accounted for by the fact that missing scores for a certain minimum number of tests were interpolated in making the composites, by a method that will be explained in a later section.

The schools were chosen so as to represent different types as to size and sociological conditions of environment. School No. 1 is a small, comparatively new school in the outskirts of the city, drawing some pupils from less well-graded districts. The community is largely composed of people of the artisan classes. School No. 2 is also a small school, but in an old and rather unprogressive section of the city. No. 3 is a large cosmopolitan school with a varied school population. No. 4 is also a large school and is in one of the best residence sections of the city largely composed of business and professional classes.

Table I shows the distribution of classes by schools, rooms, and grades.

TABLE 1

DISTRIBUTION OF CLASSES BY SCHOOLS, ROOMS, AND GRADES

School	No. of Rooms	Number of Classes			
		Grade VI	Grade VII	Grade VIII	Total
No. 1.....	5	2	3	2	7
No. 2.....	4	2	3	2	7
No. 3.....	7	3	3	2	8
No. 4.....	8	3	3	3	9
All.....	24	10	12	9	31

2. TESTS USED

The tests used were:

Addition, Woody Scale, Series A
Subtraction, Woody Scale, Series A
Multiplication, Woody Scale, Series A
Division, Woody Scale, Series A
Problems I
Problems II
Problems III
Trabue Completion-Test Language Scale B
Trabue Completion-Test Language Scale C
Trabue Completion-Test Language Scale D
Trabue Completion-Test Language Scale E
Composition I. A Letter Applying for a Job.
Composition II. On the topic, "The Study I Like Best and Why, and the Study I Like Least and Why."
Paragraph Reading, Thorndike Reading Scale Alpha 2
Spelling, last 20 words in column "U" of Ayres' Measuring Scale for Ability in Spelling
Opposites, A1
Opposites, A2
Opposites, A3
Opposites, A4
Whole-Part, B1
Adjective-Substantive, B2
Verb-Object, C1
Species-Genus, C2
Part-Whole, C3
Mixed Relations, D1
Mixed Relations, D2
Directions, X
Directions, VI
Directions, VII
Visual Vocabulary, VIII
Visual Vocabulary, XIII

Specific references to the sources of those tests which are readily available are given in the appended bibliography, as follows:

Woody Arithmetic Scales (27)
Trabue Language Scales (25)
Thorndike Reading Scale Alpha 2 (20)
Ayres' Measuring Scale for Ability in Spelling (1)
Directions, X (24)
Directions, VI (24)

The other tests are reproduced below.[1]

A-1.

Write your name here..... Write your age here..years..months.
Write the date here.....

Write after each word on this page its *opposite* as shown in the first five. If you do not know the opposite of a word go on to the next word.

good—*bad*
day—*night*
up—*down*
friend—*enemy*
true—*false*

high	after
summer	above
long	sick
soft	slow
white	large
far	rich
up	dark
smooth	front
lost	love
wet	tall
high	open
dirty	summer
east	new
day	come
yes	male
wrong	to reveal
empty	level
top	past
north	common
sour	foreign
out	sane
weak	country
good	rapid

[1]The writer is indebted to Professor Thorndike for these tests.

A-2.

Write your name here..... Write your age here..years..months.
Write the date here.....

Write after each word on this page its *opposite* as shown in the first five. If you do not know the opposite of a word go on to the next word.

good—*bad*
day—*night*
up—*down*
friend—*enemy*
true—*false*

early	extravagant
dead	obscure
hot	opaque
asleep	convex
serious	heterodox
grand	collect
to win	prompt
to respect	strong
clean	to lead
few	courteous
alike	tragic
deep	diminish
hiss	erroneous
encourage	fact
coarse	stale
melt	genuine
double	broken
dwarf	permit
plural	descend
valley	expensive
to benefit	divine
future	all
spend	from

A–3.

Write your name here..... Write your age here..years..months.
Write the date here.....

Write after each word on this page its *opposite* as shown in the first five. If you do not know the opposite of a word go on to the next word.

good—*bad*
day—*night*
up—*down*
friend—*enemy*
true—*false*

frequently
to lack
apart
stormy
motion
forcible
straight
to hold
loquacious
forbid
always
grief
sickly
negative
soothing
free
superior
separate
noise
city
attractive
stupid
shallow

some
savage
different
to marry
every
masculine
obnoxious
enlarge
joy
scowl
sacred
inhale
aristocratic
help
foolish
pay
wild
precise
dwindle
drunk
fluctuating
first
horizontal

A-4.

Write your name here..... Write your age here..years..months.
Write the date here.....

Write after each word on this page its *opposite* as shown in the first five. If you do not know the opposite of a word go on to the next word.

good—*bad*
day—*night*
up—*down*
friend—*enemy*
true—*false*
after
to float
rough
to bless
to take
exciting
clumsy
unless
tender
public
raise

remember
increase
preserve
debit
if
vertical
ignorant
rude
simple
deceitful
stingy

impoverish
cruel
generous
ancient
silly
multiply
desist
survive
proficient
hindrance
strength
innocent

permanent
over
to degrade
weary
to spend
part
together
gradual
victorious
optimist
laugh
numerous

B-1.

Write your name here.....
Write your age here..years..months
Write the date here.....

Write words that fit the words in this column in the way shown in the first three.

face—*nose*
tree—*branch*
store—*counter*
apple
clock
knife
book
hat
pencil
hand
dog
oyster
church
chair
bird
banana
shoe
train
finger
house
coat
cart
face

B-2.

Write your name here.....
Write your age here..years..months
Write the date here.....

Write words that fit the words in this column in the way shown in the first three.

wet—*water*
red—*rose*
sharp—*knife*
sharp
hot
dusty
raw
deep
ripe
funny
tall
stormy
new
hilly
strong
muddy
pretty
noisy
white
steep
round
smoky
curly

C-1.

Write your name here..... Write your age here..years..months.
Write the date here.....

After each word printed below you are to write some word, according to the further directions. Write plainly. If you cannot think of the right word, *go ahead to the next.*

Write words that fit the words in this column, in the way shown in the first three.

drink—*water*
ask—*questions*
subtract—*numbers*

bake	light
spill	sail
kiss	spin
polish	lock
sweep	wash
fill	pump
sharpen	learn
write	open
chew	eat
drive	climb
read	lend
tear	smoke
throw	singe
paint	dig
mail	sift

C-2.

Write your name here..... Write your age here..years..months.
Write the date here.....

After each word printed below you are to write some word, according to the further directions. Write plainly. If you cannot think of the right word, *go ahead to the next.*

Write words that tell what sort of a thing each thing named is, as shown in the first three.

lily—*flower*
blue—*color*
diamond—*jewel*

penny	iron
dictionary	green
cabbage	42
Rhine	parlor
murder	ruby
dog	orange

sparrow	cat
London	September
foot-ball	dime
rose	elm
diphtheria	France
robin	skirt
Pacific	a rattle
cod	Christmas
baseball	muslin

C-3.

Write your name here..... Write your age here..years..months.
Write the date here.....

After each word printed below you are to write some word, according to the further directions. Write plainly. If you cannot think of the right word, *go ahead to the next.*

Write words that fit the words in this column, in the way shown in the first three.

sleeve—*coat*
nose—*face*
roof—*house*

elbow	sleeve
hinge	brick
page	deck
finger	France
wing	pint
morning	fin
blade	steeple
mattress	month
chimney	hub
cent	chin

D-1.

Write your name here..... Write your age here..years..months.
Write the date here.....

Write in each line a fourth word that fits the third word in that line in the way that the second word fits the first, as shown in the first three lines.

color—*red*	name—*John*		
page—*book*	handle—*knife*		
fire—*burns*	soldiers—*fight*		
good—bad	long—	deep—valley	high—
eagle—bird	shark—	growls—dog	roars—
eat—bread	drink—	brick—wall	page—

fruit—orange	vegetable—	lathe—machine	hammer—
sit—chair	sleep—	pencil—lead	book—
double—two	triple—	high—low	up—
England—London	France—	sheep—lamb	dog—
chew—teeth	smell—	Thursday-Friday	June—
pen—write	knife—	build—house	paint—
water—wet	fire—	one—single	two—
laborer—works	soldier—	eye—see	ear—
come—came	go—	Monday-Tuesday	April—
north—south	far—	do—did	see—
mend—clothes	bake—	bird—sings	dog—
lily—flower	oak—	hour—minute	minute—
ton—pound	pound—	straw—hat	leather—
elbow—arm	chin—	cloud—rain	sun—
pea—pod	nut—	hammer—tool	dictionary—
past—present	present—	uncle—aunt	brother—
bell—rings	clock—	dog—puppy	cat—

D-2.

Write your name here..... Write your age here..years..months.
Write the date here.....

Write in each line a fourth word that fits the third word in that line in the way that the second word fits the first, as shown in the first three lines.

color—*red*	name—*John*		
page—*book*	handle—*knife*		
fire—*burns*	soldiers—*fight*		
he—him	she—	dusty—dry	muddy—
boat—water	train—	floats—raft	sails—
crawl—snake	swim—	deck—ship	mattress—
horse—colt	cow—	iron—metal	Latin—
nose—face	toe—	cat—kitten	hen—
bad—worse	good—	dark—light	hot—
hungry—food	thirsty—	month-December	holiday
hat—head	glove—	spin—top	sharpen—
ship—captain	army—	second—minute	minute—
man—woman	boy—	pencil—lead	book—
axe—cuts	pin—	little—less	much—
yard—foot	foot—	wash—face	sweep—
early—late	new—	house—room	book—
shoot—gun	smoke—	sky—blue	grass—
Atlantic—ocean	Mississippi—	swim—water	fly—
has—had	is—	once—one	twice—
month—year	day—	cat—fur	bird—
my—mine	they—	pan—tin	table—
room—ceiling	house—	buy—sell	come—
baby—cries	bird—	oyster—shell	banana—

VII

DO WHAT IT SAYS TO DO

21. Show by a cross which costs most:
an orange
a suit of clothes
a pair of skates
a pound of sugar

22. Show by a cross which tastes best:
dirty paper
coal dust
roast beef
sour milk

23. How many legs has a horse? Write the answer.

24. How many more legs has a horse than a boy? Write the answer.

25. Show by a cross which you would like best to own:
a toy boat
a gold ring
a thousand dollars
a sharp knife

26. Show by a cross which you would enjoy most:
being hit with a club
having your shoulder pinched
receiving a gift of money
losing many toys

27. Show by a cross the word that means pleasant:
entrance
entreat
entangle
enjoyable

28. Show by a cross the words that tell something a boy ought to do:
perform miracles
overeat habitually
study his lessons
strain his eyes

29. Show by a cross which is hardest to do:
To run a mile in a second
To stay awake all night
To pay attention in school
To play out-doors

30. Show by a cross the least dangerous thing:
To sit in front of a trolley car
To sit in the snow all night
To play with a loaded gun
To sit still in front of a mirror

31. Show by a cross each word that means unhappy.

grotesque
wretched
sincere
miserable
sad
notable
sane
joyless
grateful
uncomfortable

32. Show by a cross each word that means *to make clear* or something like *to make clear:*

execrate
elevate
elucidate
traduce
explain
clarify
satisfy
expound
extort
antipathy

VIII

Write a letter *b* under every word that means some part of the body. Write a letter *t* under every word that is the name of a tool. Write a letter *s* under every word that means something to do with the sea or ships. Write a letter *h* under every word that means some part of a house.

REMEMBER—*b*, for parts of the body
t, for tools
s, for words about the sea and ships
h, for parts of a house

arms, ear, wave, door, eye, ship, hall, saw, bone, sail, hammer, wall, ocean, face, deck, window, axe, float, mast, chamber, foam, file, canoe, billow, cellar, knife, harbor, elbow, coast, attic, brain, bosom, basement, breast, launch, artery, chisel, shin, ceiling, brace, tongue, helm, garret, porch, gulf, kitchen, hatchet, fleet, balcony, gimlet, channel, casement, cargo, entry, cruise, embark, chimney, thigh, tiller, awl, alcove, ell, keel, piazza, cleaver, spine, screw-driver, dormer, bevel, cartilage, rudder, corridor, eaves, ratchet,

kidney, hurricane, gable, lathe, cerebellum, lagoon, plane, mallet, leeward, hearth, vise, cranium, nautical, wainscot, adze, ligament, cornea, augur, navigable, patella, starboard, lymph, mariner, peritoneum, hull, jetty.

XIII

Look at each word. Think what it means. If it means a color like *red* or *blue*, write a letter *c* under it. If it means something about a number, like *six*, *all*, *half*, *many*, write *n* under it. If it means something about time, like *now*, *late*, *when*, *never*, write *t* under it. If it means something about direction or location, if it is a word like *east*, *north*, *up*, *down*, *above*, *behind*, write *d* under it.

REMEMBER—*c*, for words about color, like *red*, *blue*, *green*
n, for words about number, like *more*, *five*, *many*
t, for words about time, like *then*, *often*, *early*
d, for words about direction and location, like *front*, *east*, *here*

last, west, here, half, early, brown, many, below, there, month, across, year, noon, back, forty, gray, first, right, pair, left, green, morning, whole, pink, while, where, afternoon, minute, gold, edge, soon, outside, some, more, evening, plum, inside, beneath, odd, even, every, yellow, monday, toward, dozen, june, group, future, buff, from, count, nine, forenoon, purple, middle, plural, tuesday, instant, couple, score, crimson, april, each, during, interior, forward, wednesday, never, december, singular, center, hazel, outward, several, scarlet, violet, double, frequent, chestnut, august, ruby, either, rarely, recent, overhead, million, tan, exterior, quartet, seldom, tawny, olive, january, margin, numerous, immediate, bounding, february, encircling, september, lilac, numeral, crowd, gradual, roan, quadruple, opposite, minority, november, russet, diagonal, october, faun, formerly, garnet, overhanging, multitudinous, lasting, meantime, duplicate, continue, intervening, perpendicular, henceforth, elevated, lavender, forthwith, sextet incessant, azure, surmounting, majority,

previous, interim, thereabouts, plurality, narrow, perennial, frontal, treble, repeatedly, constantly, carmine, neighboring, parallel, contemporary, vertical, mauve, decade, emerald, magenta, octile, ecru, contiguous, quintet, ensuing, horizontal, evanescent, sepia, conterminous, multiple, integer, vermillion, eternal, turquoise, transverse, century, sexagesimal, ephemeral.

Problems—I

Begin with No. 1 and see if you can solve all ten correctly.

1. How much change should I expect from $5.00, after paying for 5 pounds of coffee at 38 cents a pound?
2. A baseball team played 160 games during the season and won 100 of them. What part of the whole number of games did the team win?
3. If $1,991 a day is paid to 724 men who each earn the same wages, how much does each man receive?
4. The children of a school made badges. Five hundred and fifty were needed. In 4 days grades 3 to 7 made 20, 25, 63, 132 and 144 badges. The eighth grade agreed to make the rest. How many did the eighth grade agree to make?
5. A man has a salary of $125 a month. He saves 20 per cent of his salary. How much will he save in a year?
6. A store takes in the following sums: $1,250.50, $300, $175, $16.25, $120.50, $32.75, $68.50. It pays out: $600, $360, $166.67, $44.33, $240. How much remains after payments are made?
7. Mr. Marshall receives a salary of $2500 a year. His rent costs him one-fifth of this and his other expenses are $1,500. He saves the rest. What per cent of his salary does he save?
8. A grocer had a tank holding 44⅚ gallons of oil. One day he drew out 15¾ gallons and the next day 9⅛ gallons. How many gallons were left in the tank?
9. How much will Mr. Fox receive for 8¾ dozen pencils at the rate of 6 for 25c?
10. John is 4 ft. 9 in. How tall will he be in two years if he grows 3¼ per cent of his present height the first year, and 3¼ per cent of his height a year from now the second year?

Problems—II

Find how long Mary was allowed to play on each of these days.

1. Monday. It is 4.10 P. M. Mother says to Mary, "You may play till quarter past five."

2. Tuesday. It is 4.20 P. M. Supper is at 6 o'clock. Mother says, "You may play half the time from now till supper time."

3. Wednesday. It is 4.05 P. M. Mother says, "If you will help me for half an hour now, and for 10 minutes before supper you may play the rest of the afternoon."

4. Thursday. Mother says, "You may play 20 minutes and 2½ minutes more for every piece you iron." Mary irons 28 pieces.

5. Friday. Mother says, "You may play 2 minutes for every 3 problems you solve, and 5 minutes more for every problem you solve correctly." Mary solved 15 and had all but one right.

6. Saturday. Mother says, "You may play 2 hours for nothing. Also I will allow you 10 minutes for every problem you solve correctly, but will take off 10 minutes for every problem that is wrong. Also you may play 1½ minutes for every minute you help me by minding your little brother." Mary did 10 problems, and had only one of them wrong. She minded the baby for an hour and a quarter.

Problems—III

See how many of these problems you can do correctly. All the problems ask the same question, "How many minutes is it from the time John begins to pump until the tank is filled?" The tank holds 120 gallons and is supposed always to be empty when John begins to work.

1. John pumps 2 minutes before any water reaches the tank. Then he pumps water into it at the rate of 3 gallons a minute until the tank is full.

2. John pumps 1½ min. before any water reaches the tank. Then he pumps water into it at the rate of 3 gallons a minute, for 20 min. Then he pumps at the rate of 2 gallons a minute until it is full.

3. John pumps 1 min. before any water reaches the tank. Then he pumps water into it at the rate of 24 gallons in 10 minutes until the tank is full.

4. John pumps 2 min. before any water reaches the tank. Then he pumps for 15 min. at the rate of 3 gallons per minute. Then Dick helps him and they pump at the rate of 5 gallons per minute until the tank is full.

5. John pumps 1¾ min. before any water reaches the tank. Then he pumps for 10 min. at the rate of 2.7 gallons per minute. Then the pump breaks and he spends 8 min. mending it. Then he pumps at the rate of 3.1 gallons per minute until the tank is full.

6. John pumps for 1 min. 50 sec. before any water reaches the tank. Then he pumps at the rate of 3.6 gallons per minute for 10 min., then rests 5 min., then pumps 3.6 gal. per minute for 10 minutes, then rests 5 min., then pumps 3.6 gal. per min. for 10 minutes, then rests 5 min., and so on until the tank is filled.

The selection of the tests used was made with the following aims in mind:

a. To make as thorough a testing as time would permit.

b. To have the two general fields of Arithmetic and English well represented.

c. To use the best available tests in school subjects, suited to the purpose of this study.

The Woody Arithmetic Scales furnish a real test for the sixth, seventh, and eighth grades, present a variety of types of operations, and are a test of power rather than a speed test.

The Problem Tests were known to be unfamiliar to the pupils. The form of II and III is such as to test for a minimum of control of vernacular and a maximum of control of operations and mathematical reasoning.

The Trabue Completion Tests represent the scientifically derived results of much careful testing. Dr. Trabue says of them, "It will be found that ability to complete these sentences successfully is very closely related to what is usually called 'language ability'." (25, p. 1)

Thorndike Reading Scale Alpha 2 is a scientifically sound scale for the measurement of paragraph reading. The only limitation is that there are not other comparable scales available.

Composition. A letter was chosen as one form of composition because it is commonly taught in schools and is the most ordinary form of writing done outside of school. The topic, "The Study I Like Best and Why, and the Study I Like Least and Why,"[1] was chosen as furnishing a theme on which every child had something to say and as giving scope for the more capable pupils.

---

[1]This topic is suggested to teachers as of particular value from the standpoint of throwing light on the "content of children's minds" with respect to their studies. There is, no doubt, an advantage in having the test given by an outsider as the pupils tested gave evidence of more frankness than might be expected from them if writing for the teacher to read.

The Ayres Spelling Scale furnishes a comparatively large list from which to choose, and hence one less likely to have been drilled upon. Column "U" was chosen as representing a degree of difficulty not altogether beyond reasonable attainment by the sixth grade and yet such as to test the eighth.

d. To supplement these with a varied selection of psychological tests. Being tests involving knowledge of verbal relations, they serve to extend the scope of the English group.

Opposites tests have been often shown to be very satisfactory tests as a measure of the control of the vernacular.

Visual Vocabulary tests supplement testing of paragraph reading. They are found to rank high as measures of "language power."

Directions tests furnish a very real test of power to get meaning from a printed page.

The other controlled association tests used had no particular claim except as supplementary to the others.

3. Administration of the Tests

Uniformity of Procedure.

The aim in this study being a comparison of achievements in the different grades tested, it was of primary importance that the procedure in giving the tests be uniform throughout. Considering the number of tests used, the number of schools, rooms, and classes concerned, involving numerous repetitions, the necessity for rigid control of conditions during the testing is obvious. To this end the following procedure was maintained throughout.

a. Every test exercise was started by the writer.

b. The writer remained in the room throughout the first test. Thereafter, except in the case of the short tests, he went to another room leaving an assistant to receive the papers and record the time on each.

c. Either the writer or an assistant was in the room throughout each test.

d. The teacher was regularly present, but ordinarily engaged in other work at the rear or side of the room. The very best of cooperation on the part of the teachers made possible very rigid control of the conditions in the room so far as any possible influence on the work of the pupils by the teachers was concerned. It was thoroughly understood from the beginning that all directions were to come from the testers. On the other hand the presence of the teacher in the room was a real,

though unadvertised, help in insuring independent work on the part of each pupil. It is a pleasure, incidentally, to record the fact that the pupils were remarkably free from any tendency to depend on others for aid.

e. At the very start of the testing in each room the mechanics of procedure in the way of provision of pencils in good condition, clearing of desks, passing of papers, putting on of headings, signal for beginning and procedure on finishing the test, were fixed by rigid adherence to a pre-arranged plan, with modifications only to suit the needs of different tests.

f. Preliminary to the beginning of each test specific directions to be given the pupils were worked out. These were adhered to carefully throughout. Any slight changes in general procedure were made at the beginning of a test, the aim being to maintain the very same procedure throughout all the rooms for a given test.

g. The same order of tests was followed in all rooms with only a few slight changes made necessary by the time schedule.

h. Special care was taken that those tests which by the nature of their content might easily become matter of common knowledge among the pupils of a school, such as spelling for example, were given during the same day's session.

i. At the cost of considerable time no test was begun at such a time as would not insure ample opportunity for all to finish before an intermission as for recess or the noon hour, except in a very few cases where the custom of the room was well established to finish any given task even after other classes had been dismissed.

Special Features.

Spelling. Twenty sentences were framed, each containing one of the twenty words of the list, care being taken to use such a construction that the word could not be mistaken. These sentences were used throughout the testing. Previous to the test period papers with the regular heading and numbered from 1 to 20 at the left-hand margin, had been prepared by the pupils under the direction of the teacher. The following directions were given:

"I shall pronounce a word; then use it in a sentence; then pronounce it once more. Don't write anything until I have done these three things.

"When you are all through with one word I shall say, 'Next.' If anyone is not ready, raise your hand. When all are ready I shall give the number of the next word. Make sure that when I say 'two' you are writing the second word. If you can't write some word leave its place blank.

"Write plainly. What I can't read is called wrong, and there is some writing I can't read."

Reading. The Thorndike Reading Scale Alpha 2 is printed on two large sheets. The exercise headed Set IV, the last one on the first sheet, is repeated as the first exercise on the second sheet. In order to make sure that the pupils would not spend time on this second copy of Set IV, that part of the second sheet was cut off.

Headings. Contrary to the usual practice the headings were put on the back side of the test papers. This insures that on the signal to begin all start work on the test material, and eliminates the variation in the length of time required to write the heading, as a source of error.

Distribution of Papers. All test papers were distributed, always with the printed side down and the top end away from the pupil, by the testers. This makes for a saving of time and regularity of procedure.

Timing. As the time element is not made a feature of this study only a brief statement will be made regarding the timing of the tests. The following directions were given:

"You will all begin at the same time. When you have done everything as well as you can, bring your paper to the desk."

The time for each paper was recorded in minutes and seconds at the moment it was laid on the timer's desk—the teacher's desk in the front of the room. In case of the very short tests two timers were used in order to avoid delay in recording. In a group test without a time limit there is an unavoidable source of error in getting each individual's time recorded. It was decided this could be reduced to the minimum by taking the time at which the paper was put on the timer's desk. This method works a slight injustice to those in the rear of the room, but results in no constant error.

The Tests as a Real Measure of Capacity.

The statement was made at the beginning of this chapter that much of whatever value this study might have would hinge upon the extent to which each pupil was given an opportunity to show his real capacity in each test. To this end no test was given as a speed test; each pupil was given as much time as he wished. The assumption here is that in this way we get a much more accurate measure of what each pupil actually is capable of doing in his work from day to day. Our belief is that very little of the

work on which his grade standing is based counts speed as an important factor. Ordinarily there is ample time for the tasks required. We do not raise the question as to whether this is as it should be. Our purpose is to compare the attainments of different school grades as they exist.

However, while it was the intention to allow the pupils all the time they needed, some stimulus against the waste of time was necessary in the interest of economy. This was provided by the fact that the pupils knew their time was recorded. There certainly was some tendency to undue haste especially in the shorter tests. Except for a few of these short tests (the B–C group), however, the writer is confident this was a negligible factor. Early in the testing the practice of turning to other work as soon as a test was finished was well established among the pupils. As a result little was known as to who was not through with the test and hence no particular social disapproval was attached to slowness. The further fact that all the pupils had been subjected to testing under controlled conditions made the experience not unlike regular school work.

In the case of the longer tests, as for example Alpha 2, in order not to interfere too much with the regular schedule, pupils who had not finished within a reasonable time were taken to an unoccupied room and there permitted to finish.

Another factor making for an adequate measure of the attainments of the pupils was the large number of occasions on which the testing was done. The number of half-days during which one or more tests were given was 10, 11, 15, and 17 respectively for the different schools. This fact tends to offset the error which comes from using a single day's results as the measure of a pupil's standing. None of us care to be judged on the work of a single day, and with good reason. The following statement by Professor Hollingworth (12) brings out the significance of this point: "The momentary ability revealed in initial trials, or even in the first half dozen trials in a given set of tests might well be expected to show only low degrees of correlation. These would not be measures of ultimate capacity, but would be highly determined by previous practice, chance variability, momentary attitude and initial method of attack. They would, in short, be samplings only of momentary ability, not of final capacity."

4. Scoring the Tests

Consistency in the scoring of test papers is perhaps only second in importance to uniformity of procedure in giving the tests. This calls for particular care when the number of test papers in a given test is so large as to make impossible their scoring at one sitting, or indeed at a half dozen, but requires many sessions even though long. Furthermore, the large number of papers makes practically impossible the scoring by one person. Hence the need of careful supervision and accurate recording of method from the beginning to be followed throughout a given test. Of course the difficulty in scoring consistently is far less in the case of tests of the right or wrong sort, than in the case of those calling for partial credit and involving judgment; but even in the former the possibility of error in a large number of tests is considerable.

All the test papers in this study were scored by the writer or under his direct supervision. By far the larger part of the scoring was done by three readers, three others having done a small part.

The purpose being to get results comparable with each other rather than with previously obtained data, no particular effort was made to use methods absolutely identical with those used elsewhere. This was at times sacrificed for the much more important consideration of consistency throughout this study. However, for the benefit of any who may wish to compare other results with those here presented the following statement is made of the method of scoring.

Woody Arithmetic Scales

Score: Number of examples correct

Problems

Score: Number of points on basis of

3 for full credit

2, 1 } for partial credit

0 for no credit

The answer recorded in the place assigned was the criterion used.

Trabue

Score: Number of points on basis of

2 for full credit

1 for partial credit

0 for no credit

COMPOSITION

a. The composition on the topic, "The Study I Like Best and Why and the Study I Like Least and Why" was scored with the use of the Thorndike Extension of the Hillegas Scale. (23)

b. The letters were scored by means of an improvised 'scale.' Briefly the method of preparing this 'scale' was as follows. Forty papers were selected from the lot, representing in the judgment of the writer the whole range of ability from the poorest to the very best. It was not assumed that this number of different qualities could be detected. This large number was selected with the purpose of getting all grades of quality. Five competent judges then ranked the papers according to the directions below:

"Please rank the papers in eight groups of five each calling Group I the poorest and Group VIII the best.

"Enter the number of each paper (appearing in the upper right hand corner) on the appropriate line below.

"Note that you are not asked to rank the papers within each group.

"Assuming arbitrarily the value 10 for Group I and the value 80 for Group VIII, assign intervening values to the other groups. Enter these values at the left of the number of the group."

From the results of these rankings a selection of eight papers was made, those being selected on which there was greatest agreement among the judges as to the group in which each belonged. Of two or three papers equally placed, that one was chosen which seemed to best fit into the whole. The amount of agreement among the judges is indicated below:

Number placed in same group	by	Number of judges
2		5
10		4
10		3

On the combined judgments of the five readers the values from 10 to 80 by equal intervals of 10 were given to the compositions in the eight groups from the poorest to the best. This is a rough evaluation and lays no claim to scientific accuracy. But the results justify the use of the 'scale' for the purpose of this study. The aim was to overcome the difficulty which all readers have who try to use the Hillegas Scale to judge compositions in the form of letters. The arbitrary method of assigning values to the letters composing the 'scale' was with the purpose of getting values roughly comparable with those on the Hillegas Scale.

In all the work of scoring the compositions the reader was ignorant of the grade in which a given composition was written, this being very essential if a comparison of results by grades is to be made.

The scoring was done by three readers, each paper being scored by each reader independently of the others. This gave three scores for each paper or six scores for each pupil. The final score for each pupil was obtained by taking the median of these six measures.

THORNDIKE READING SCALE ALPHA 2

Score: Number of correct responses. The key published by the author (20) was used, with additions as necessary.

SPELLING

Score: Number of words spelled correctly

OPPOSITES

Score: 3 for full credit
2 } for partial credit
1 }
0 for no credit

All the available keys were used. However, numerous answers had never been passed upon in these keys and it became necessary to extend them greatly. This was done by using the combined judgments of from two to eleven judges. Because of the very great number of answers occurring only a very few times it was necessary in these cases to rely on the judgments of only two persons, the writer and one other in every case.

B1, B2, C1, C2, C3

Score: 1 for full credit
½ for partial credit
0 for no credit

D1, D2

Score: Number of correct responses

X

Score: A, B, D, E, F were allowed 2 each when correct. C and G were allowed 1 for each line correct.

VI

Score: Number of correct responses

VII

Score: One point was allowed for each step correct except 31 and 32. These two steps were scored as a unit. The number right minus the number wrong, counting omissions, being the score.

VIII

Score: Number of correct responses

XIII

Score: Number of correct responses

III

OVERLAPPING BY SINGLE TESTS

1. The Measure of Overlapping

The measure of overlapping in this study is the per cent of the pupils in any grade who equal or exceed the median score of the next grade or the second grade above; or the per cent of the pupils in any grade who go below the median score of the next grade or the second grade below. The former will be referred to as overlapping "upward" and the latter as overlapping "downward."

The value of this measure of overlapping is shown in the following statements by Professor Thorndike, and by his accompanying diagrams. (21) "The great advantage gained by comparing groups by the per cent of one group reaching or exceeding the point on the scale that is reached or exceeded by a given per cent of the other group is that results are mutually comparable whatever the traits may be . . . Another advantage lies in the fact that this percentile comparison reminds one constantly of the overlapping of the two groups, when such exists."

The distributions on which the per cents of overlapping are calculated will be found in the Appendix. Table II shows the amount of overlapping in 21 tests, of the sixth grade on the eighth, the sixth on the seventh, and the seventh on the eighth, upward; and downward, the eighth on the sixth, the eighth on the seventh, and the seventh on the sixth.

TABLE II

OVERLAPPING OF GRADES BY SINGLE TESTS

Per cent in each grade who equal or exceed the median of the other grade; upward. Per cent in each grade who go below the median of the other grade; downward

	Upward			Downward			
Tests	VI on VIII	VI on VII	VII on VIII	VIII on VI	VIII on VII	VII on VI	Average of the six measures
Addition	24.19	35.92	38.71	25.90	36.15	37.39	33.04
Subtraction	22.30	25.66	44.22	19.90	42.09	25.77	29.99
Multiplication	23.42	33.23	40.90	26.53	39.18	34.92	33.03
Division	28.27	35.36	43.57	26.76	38.73	37.43	35.02
Problems I	7.14	13.31	32.77	9.63	33.49	18.45	19.13
Problems II	10.64	16.95	36.06	14.04	37.19	25.29	23.36
Trabue B	26.61	34.54	44.29	30.01	42.86	38.12	36.07
Trabue C	21.68	38.49	34.65	25.28	33.32	40.05	32.25
Trabue D	23.31	38.58	36.95	30.92	38.34	41.32	34.90
Trabue E	13.99	36.86	25.91	19.74	26.97	37.16	26.77
Alpha 2	11.61	28.02	31.99	12.73	33.00	26.49	23.97
Spelling	7.37	21.09	30.26	7.39	26.34	22.14	19.10
Opposites A2	8.76	19.37	36.57	14.08	35.44	23.65	22.98
Opposites A3	13.33	31.80	28.44	14.16	26.41	32.82	24.49
Opposites A4	10.05	22.18	37.77	11.15	34.11	27.50	22.79
Part-Whole C3	35.18	45.59	35.42	34.85	37.82	45.77	39.11
Mixed Relations D1	42.53	53.50	43.95	43.87	41.19	52.65	46.28
Mixed Relations D2	38.75	48.67	39.50	34.60	35.84	49.07	41.07
Directions VII	21.23	37.38	35.10	27.88	38.80	38.82	33.20
Visual Vocabulary VIII	10.64	24.22	24.06	9.20	28.34	24.69	20.19
Visual Vocabulary XIII	11.06	26.53	29.94	11.41	28.29	23.40	21.77
Average	19.62	31.77	35.48	21.43	34.95	33.47	29.45
Range	7.14 to 42.53	13.31 to 53.50	24.06 to 44.29	7.39 to 43.87	26.34 to 42.86	18.45 to 52.65	19.10 to 46.28

The table reads: In addition 24.19 per cent of the pupils in the sixth grade equal or exceed the median of the eighth grade; 38.71 per cent of the pupils in the seventh grade equal or exceed the median of the eighth grade; 25.90 per cent of the pupils in the eighth grade go below the median of the sixth grade; and 36.15 per cent of the pupils in the eighth grade go below the median of the seventh grade.

2. The Elimination of Some Tests

The fact that there are figures for only 21 tests calls for explanation. It is the purpose to present the results for only those tests which proved to be fair measures for the three grades used, on the basis of the distributions of the scores for each grade. Some tests proved too easy with consequent piling up of undistributed scores at the upper end. The higher ranges of ability were not measured. The way up was closed, so to speak. This naturally works to the greatest disadvantage of the eighth grade, to less for the seventh, and least for the sixth. The result is an unduly high per cent of overlapping upward. The closeness with which the grade medians approximate the total possible score is a rough measure of this exaggeration of the overlapping figures. For example, the median score of the sixth grade may be very near the maximum possible score, beyond which obviously none in the eighth grade can go. Table III has been prepared to show which tests are affected in this way. This table gives the median score for each grade in each test and the maximum possible score. With the purpose, then, of ruling out those tests which by their very nature, as shown in the distribution tables, make a high per cent of overlapping inevitable, the tests A1, B1, B2, C1, C2, X, VI are omitted from Table II. Problems III proved much too difficult with undistributed scores at the lower end, and hence is not included.[1] The two composition tests, the scores for which were not distributed singly, do not appear in the table. Composition is treated only as a composite.

It should be borne in mind in connection with the treatment in Chapter IV that the inclusion of the first eight tests mentioned above would have raised the average per cents of overlapping to higher figures. Hence any showing based upon the magnitude of the per cents of overlapping by single tests would have been magnified had these tests been included. There is no absolute line of demarcation between the tests that should and the tests that should not be used. The error, for the validity of the argument which follows in Chapter IV, could not have been in the direction of eliminating too many tests.

---

[1]A table of per cents of overlapping for these tests is given in the Appendix.

TABLE III

MEDIAN SCORES IN EACH TEST FOR EACH GRADE, WITH MAXIMUM POSSIBLE SCORES

Tests	Grades VI	VII	VIII	Maximum possible score
Addition	29.30	30.89	32.35	38
Subtraction	25.74	28.37	28.87	35
Multiplication	27.84	29.65	30.69	39
Division	24.54	26 06	26.85	36
Problems I	9.11	15.36	18.35	30
Problems II	2.28	5.06	6.54	18
Trabue B	13.37	14.13	14.57	20
Trabue C	12.94	13.57	14.65	20
Trabue D	14.20	14.61	15.32	20
Trabue E	13.30	13.93	15.50	20
Composition	37.66	42.89	49.27	87.5
Alpha 2	22.61	25.37	27.08	38
Spelling	12.73	17.15	18.97	20
A 1	122.69	126.19	128.77	138
A 2	53.71	66.11	73.44	138
A 3	49.52	56.50	68.07	138
A 4	48.45	65.08	73.37	138
B 1	16.50	18.29	17.57	20
B 2	*	*	*	20
C 1	29.28	29.62	*	30
C 2	26.27	27.56	28.31	30
C 3	15.34	15.71	16.69	20
D 1	16.12	15.76	16.85	40
D 2	17.19	17.45	22.25	40
X	14.69	16.10	16 11	18
VI	19.43	19.55	19.55	20
VII	10.17	10.59	11.33	18
VIII	70.12	78.10	84.67	100
XIII	75.38	100.63	119.35	169

3. COMPARISON OF DIFFERENT TESTS AS TO AMOUNT OF OVERLAPPING

We have been prepared to find large amounts of overlapping and we find our results no exception. Less to be expected, perhaps, is the great range among the tests. It is obvious that were one making deductions from overlapping figures from single tests

*See distribution tables in Appendix.

much would depend upon the particular tests chosen. For example, any conclusions regarding overlapping in arithmetic drawn from a per cent of 28.27 in division (VI on VIII) would be quite otherwise than if based on a per cent of 7.14 in Problems I. The possibility of error from inferences from such results as those presented in Table II is further illustrated by assuming that 13.99 per cent (Trabue E) measures the overlapping of VI on VIII in Completion-Test Language Scales rather than 26.61 per cent (Trabue B). Much would thus depend upon whether one used one or the other of these tests, though they are of demonstrated equal difficulty. "Language Scales D and E are practically of the same difficulty as Scales B and C." (25, p. 22)

In the discussion of the relative amounts of overlapping in the different tests which follows let it be borne in mind that we make no claim that these figures derived from single tests are valid measures of overlapping in the traits measured. The great range in the per cents for tests counted of equal difficulty has raised a feeling of doubt, to say the least. Later we shall consider this question specifically.

But taking these figures of Table II for just what they are, per cents of overlapping in the different tests as indicated, we shall see what an analysis of them reveals.

Tables IV, V, and VI have been prepared as an aid in answering the question, How do the tests rank in amount of overlapping as determined by the six sets of comparisons, three upward and three downward?

Table IV lists the tests in order of rank from the least to the most overlapping, on each basis.

TABLE IV

LISTS OF THE TESTS IN ORDER FROM THE LEAST TO THE MOST OVERLAPPING FOR EACH OF THE SIX COMPARISONS

Rank	VI on VIII	VI on VII	VII on VIII	VIII on VI	VIII on VII	VII on VI
1	Problems I	Problems I	VIII	Spelling	Spelling	Problems I
2	Spelling	Problems II	Trabue E	VIII	A3	Spelling
3	A2	A2	A3	Problems I	Trabue E	XIII
4	A4	Spelling	XIII	A4	XIII	A2
5		A4	Spelling	XIII	VIII	VIII
5.5	Problems II VIII					
6		VIII	A4	Alpha 2	Alpha 2	Problems II
7	XIII	Sub.	Alpha 2	Problems II	Trabue C	Sub.
8	Alpha 2	XIII	Problems I	A2	Problems I	Alpha 2
9	A3	Alpha 2	Trabue C	A3	A4	A4
10	Trabue E	A3	VII	Trabue E	A2	A3
11	VII	Mult.	C3	Sub.	D2	Mult.
12	Trabue C	Trabue B	Problems II	Trabue C	Add.	Trabue E
13	Sub.	Div.	A2	Add.	Problems II	Add.
14	Trabue D	Add.	Trabue D	Mult.	C3	Div.
15	Mult.	Trabue E	Add.	Div.	Trabue D	Trabue B
16	Add.	VII	D2	VII	Div.	VII
17	Trabue B	Trabue C	Mult.	Trabue B	VII	Trabue C
18	Div.	Trabue D	Div.	Trabue D	Mult.	Trabue D
19	C3	C3	D1	D2	D1	C3
20	D2	D2	Sub.	C3	Sub.	D2
21	D1	D1	Trabue B	D1	Trabue B	D1

Table V summarizes the facts of Table II on the basis of division into tertiles. It shows that Spelling and VIII fall in the lowest third in all six comparisons, no test falls in the middle and only one test, D1, falls in the highest third in the six comparisons. One test, XIII, falls in the lowest third in five comparisons, and three tests, A4, Problems II, and Problems I, fall in the lowest third in four comparisons. At the other extreme we have one test, D1, in the highest third in six comparisons, and so on. It appears therefore that there is most agreement among the different comparisons in placing D1 at the top of the list in amount of overlapping, and Spelling and VIII at the bottom of the list, that is, with least overlapping.

TABLE V

Showing What Tests are in the Lowest, Middle and Highest Thirds as to Amount of Overlapping in Four or More of the Six Comparisons, VI on VIII, VI on VII, etc.

	In six comparisons	In five comparisons	In four comparisons
In lowest third	Spelling VIII	XIII	A4 Problems II Problems I
In middle third	None	None	A3 Addition
In highest third	D1	Trabue B D2	Division C3 VII Trabue D

Table VI furnishes another basis for measuring the relative amount of overlapping for the different tests. A rank is given each test in each of the six comparisons. These ranks are added and another ranking made in column 8. A ranking in column 9 is made on the basis of the averages from the six measures. (See Table II, last column.) Finally these two rankings are added and the rankings of column 11 obtained, which may be taken as the order of the tests in amount of overlapping on this basis. From the combined results then, the following tests show the least amount of overlapping, roughly in the order named: Spelling, Problems I, VIII, XIII, A4, A2, Alpha 2.

TABLE VI

RANKINGS OF THE TESTS IN AMOUNT OF OVERLAPPING

Tests	1	2	3	4	5	6	7	8	9	10	11
	Comparisons						Sums of Rankings of Columns 1 to 6	Rankings of Totals in Column 7	Rankings of Averages of Table II	Sums of Rankings of Columns 8 and 9	Final Rankings of Totals in Column 10
	VI on VIII	VI on VII	VII on VIII	VIII on VI	VIII on VII	VII on VI					
Problems I..	1	1	8	3	8	1	22	2	2	4	2
Spelling.....	2	4	5	1	1	2	15	1	1	2	1
VIII.....	5.5	6	1	2	5	5	24.5	3	3	6	3
XIII.....	7	8	4	5	4	3	31	4	4	8	4
A4.....	4	5	6	4	9	9	37	5	5	10	5
A2.....	3	3	13	8	10	4	41	6	6	12	6
Alpha 2.....	8	9	7	5	6	8	43	7.5	8	15.5	7
Problems II.	5.5	2	12	7	13	6	45.5	9	7	16	8
A3.....	9	10	3	9	2	10	43	7.5	9	16.5	9
Trabue E...	10	15	2	10	3	12	52	10	10	20	10
Trabue C...	12	17	9	12	7	17	74	11	12	23	11.5
Sub.....	13	7	20	11	20	7	78	12	11	23	11.5
Mult.....	15	11	17	14	18	11	86	14.5	13	27.5	14
VII.....	11	16	10	16	17	16	86	14.5	15	29.5	15
Trabue D...	14	18	14	18	15	18	97	17	16	33	16.5
Add.....	16	14	15	13	12	13	83	13	14	27	13
Div.....	18	13	18	15	16	14	94	16	17	33	16.5
Trabue B ...	17	12	21	17	21	15	103	19	18	37	18.5
C3.....	19	19	11	20	14	19	102	18	19	37	18.5
D2.....	20	20	16	19	11	20	106	20	20	40	20
D1.....	21	21	19	21	19	21	122	21	21	42	21

4. SURFACES OF OVERLAPPING

Figs. 1 to 4 show in graphic form the overlapping of the three grades in tests Trabue B, Opposites–A2, Part-Whole–C3, and Division. These surfaces are all drawn of equal area and on the same base line, so that they are directly comparable one with another. The vertical distance represents the per cent of each grade attaining the scores indicated on the horizontal. The surfaces are for the sixth, seventh, and eighth grades from the bottom up in every case. The medians are drawn for each grade through the surfaces of the other grades so as to facilitate comparison. The measure of overlapping is represented by the portion of the surface of a given grade which extends beyond the median of the

grade with which comparison is made. For example, in Fig. 1 the area ABC, 34.54 per cent of the whole, represents the part of the sixth grade that reaches or exceeds the median of the seventh; and the area DBE, 26.61 per cent of the whole, represents the part of the sixth grade that reaches or exceeds the median of the eighth. Likewise, the area $A_1B_1C_1$, 30.01 per cent, represents the part of the eighth grade that goes below the median of the sixth; and the area $D_1B_1E_1$, 42.86 per cent, represents the part of the eighth grade that goes below the median of the seventh.

These figures give us also a representation of the form of distribution for these tests. Figs. 1, 2, and 4, for example, approximate rather closely to the "normal" curve, in contrast with Fig. 3, which shows considerable skewness toward the upper end. It was this condition still more emphasized that made advisable the elimination of some of the tests.

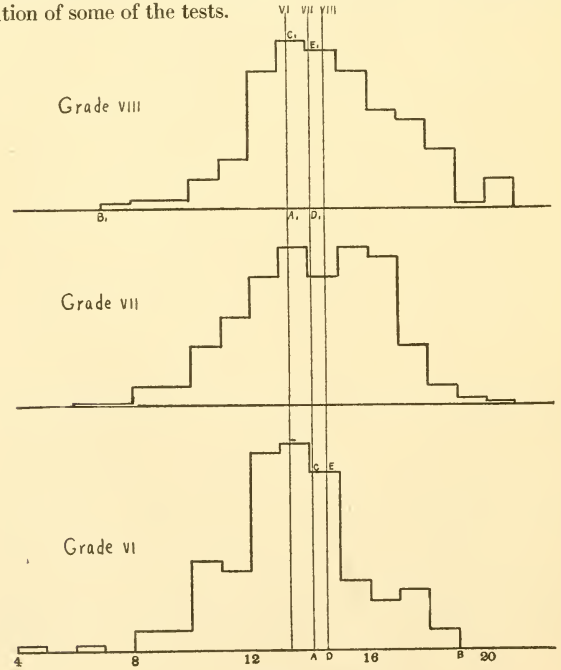

Fig. 1. Overlapping in Trabue B.

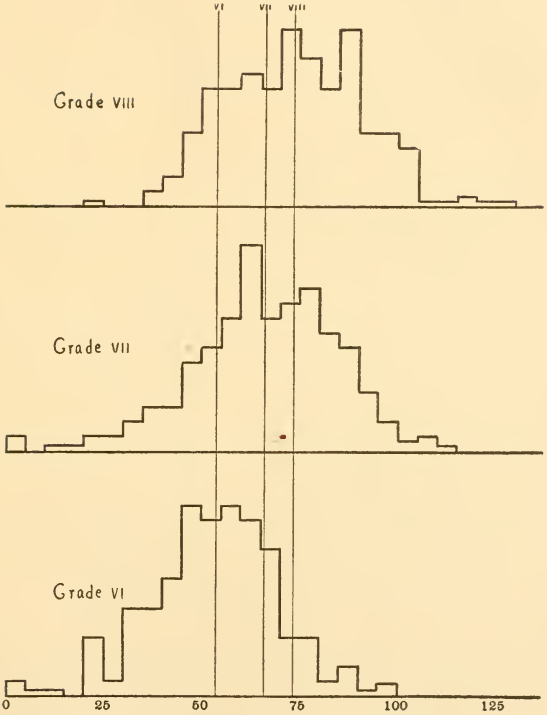

Fig. 2. Overlapping in A–2, Opposites.

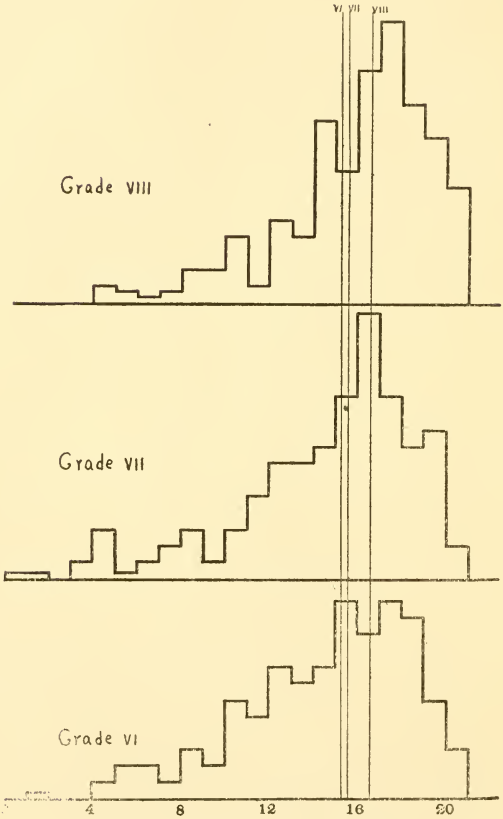

Fig. 3. Overlapping in C-3, Part-Whole.

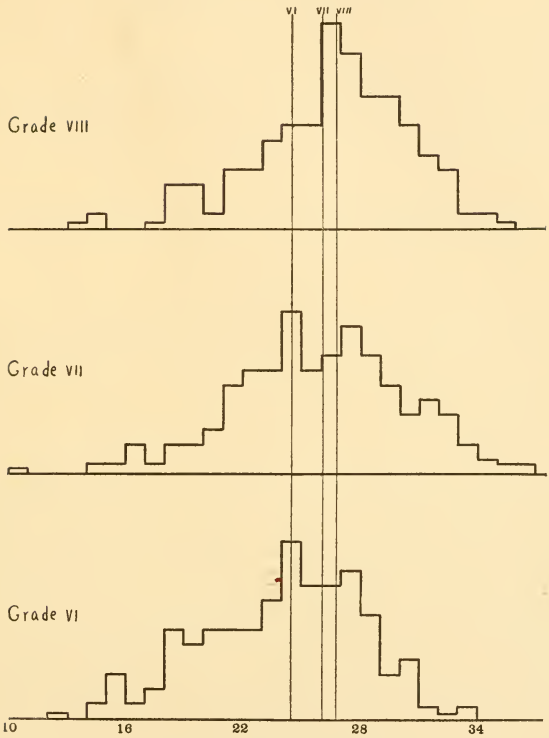

Fig. 4. Overlapping in Division.

The results presented in this chapter are in accord with those reported in the Introduction as showing great variability within the grades and consequent overlapping among the grades, when attainment is measured by single tests. Certainly measurements of this sort show large amounts of overlapping even when given under well-controlled conditions and to large numbers of pupils. We are now ready to consider the question raised at the beginning: Are measures obtained in this way valid measures of the overlapping of ability of school pupils in different grades? The next chapter will consider this problem.

IV

OVERLAPPING IN THE CASE OF COMPOSITES

1. Results

Suppose a single test be given to sixth- and eighth-grade pupils with a resulting overlapping of the sixth on the eighth of 25 per cent. Are we justified in using this figure as a measure of the overlapping of abilities in the trait measured? Only in so far, of course, as this single test is a measure of the trait. Suppose that in order to get a more valid measure of overlapping we give three more tests of the same sort and use the average of the four as the measure. This will tend to even up any extreme results from a particular test. But this average will retain any error resulting from the inadequacy of a single test as a measure of the trait in question. It does not take account of the fact that the pupil who is in the lowest quartile in one test may be in the highest quartile in the second test, and still differently placed in the third and fourth tests. In proportion as the single test is an adequate measure of a trait, a pupil's place in the distribution will be the same in the different test exercises (except in so far as particular conditions affecting the pupil at a given time influence the result). In order to give the individual pupil the benefit of more nearly adequate testing, a composite may be made of the different scores he makes in the different tests for a given trait and this used as his measure. Thus any error in his placing in the distribution resulting from only a single test tends to be offset by the other measures. The principle of procedure here is simply that up to an undefined limit the larger the number of measures, the greater the accuracy of placing the individual where he belongs in the distribution. Theoretically, then, the amount of overlapping as measured in single tests should be reduced when taken in terms of the composite.

TABLE VII

OVERLAPPING OF GRADES BY COMPOSITES

Per cent in each grade who equal or exceed the median of the other grade; upward.
Per cent in each grade who go below the median of the other grade; downward

Composites	No. of Tests	Upward VI on VIII	Upward VI on VII	Upward VII on VIII	Downward VIII on VI	Downward VIII on VII	Downward VII on VI	Average of the six measures
Composition	2	13.64	27.86	24.09	8.02	21.26	24.95	19.97
Trabue	4	14.68	32.46	31.19	21.89	31.75	36.63	28.10
Opposites	3	8.57	24.87	30.49	13.41	29.58	26.97	22.32
B–C	5	32.59	32.62	44.21	36.07	49.99	41.40	39.48
Mixed Relations	2	39.51	51.40	41.32	40.25	39.17	50.69	43.72
Directions	3	27.38	31.67	45.30	29.64	45.42	31.96	35.23
Visual Vocabulary	2	8.32	22.48	28.14	11.35	27.31	26.81	20.74
Arithmetic	6	7.08	18.13	33.21	12.00	32.68	24.28	21.23
English	22	6.97	24.86	25.83	9.68	23.18	26.12	19.44
Total	28	4.58	19.57	27.74	8.37	24.15	21.98	17.73

Table VII presents the facts of overlapping when the scores are thus combined for the different groups of tests and redistributed. Table VIII brings together the data of Tables II and VII in such a way as to permit a comparison of the results obtained by averaging the per cents of overlapping for the different tests of a group with those obtained from this method of composites. It will be seen, for example, that in the VI-on-VIII comparison the average per cent of overlapping for the four Trabue tests, 21.40, is reduced to 14.68 for the composite, a reduction of 6.72 points. Column "d" shows the difference between the two measures of overlapping, the "—" prefix indicating a reduction for the composite method. The totals for the "d" columns show that the few cases of the reverse order, an increase for the composite method, represent a very, very small portion of the total difference. Further analysis of the table reveals some interesting facts. The greatest difference in the two methods is found in the VI-on-VIII comparison, and the least in the VII-on-VIII comparison, the two in which the per cent of overlapping was, on the other hand, the least and the greatest, respectively. (See Table II.) The difference is small in the case of those composites representing a small number of tests such as Mixed Relations and Visual Vocabulary, of only two tests each. The arithmetic composite of the results from six tests is consistently lower than the average,

TABLE VIII

A Comparison of Per Cents of Overlapping in Terms of Averages of Results from Single Tests and of Composites

Composites	No. of Tests	VI on VIII			VI on VII			VII on VIII		
		Ave.	Comp.	*d*	Ave.	Comp.	*d*	Ave.	Comp.	*d*
Trabue	4	21.40	14.68	-6.72	37.12	32.46	-4.66	35.45	31.19	-4.26
Opposites	3	10.71	8.57	-2.14	24.45	24.87	0.42	32.26	30.49	-1.77
Mixed Relations	2	40.64	39.51	-1.13	51.09	51.40	0.31	41.73	41.32	-0.41
Visual Vocabulary	2	10.85	8.32	-2.53	25.38	22.48	-2.90	27.00	28 14	1.14
Arithmetic	6	19.33	7.08	-12.25	26.74	18.13	-8.61	39.37	33.21	-6.16
English*	16	19.36	6.97	-12.39	33.42	24.86	-8.56	33.30	25.83	-7.47
Total*	22	19.35	4.58	-14.77	31.59	19.57	-12.02	34.96	27.74	-7.22
Sum of *d*'s				-51.93 00.00			-36.75 0.73			-27.29 1.14

TABLE VIII—CONTINUED

Composites	No of Tests	VIII on VI			VIII on VII			VII on VI		
		Ave.	Comp.	*d*	Ave.	Comp.	*d*	Ave.	Comp.	*d*
Trabue	4	26.49	21.89	−4.60	35.37	31.75	−3.62	39.16	33.63	−2.53
Opposites	3	13.13	13.41	0.28	31.99	29.58	−2.41	27.99	26.97	−1.02
Mixed Relations	2	39.24	40.25	1.01	38.52	39.17	0.65	50.86	50.69	−0.17
Visual Vocabulary	2	10.31	11.35	1.04	28.32	27.31	−1.01	24.05	26.81	2.76
Arithmetic	6	20.46	12.00	−8.46	37.81	32.68	−5.13	29.88	24.28	−5.60
English*	16	20.96	9.68	−11.28	33.02	23.18	−9.84	34.29	26.12	−8.17
Total*	22	20.82	8.37	−12.45	34.33	24.15	−10.18	33.08	21.98	−11.10
Sum of *d*'s				−36.79			−32.19			−28.59
				2.33			0.65			2.76

*The averages for 'English' and 'Total' are for 16 and 22 tests as indicated, the per cents for the first 7 tests mentioned on p. 32 not being included. All but A1 do, however, appear in the composites. As the per cents for these tests are large (see Appendix) the differences would have been still greater had they been included. Composition counts as one test.

the difference ranging from 5.13 for VIII on VII to 12.25 for VI on VIII. So also the English composite of sixteen tests shows a reduction of from 7.47 to 12.39 below the average.

Most significant of all the figures are those for the total composite. The reduction here represents a range from 7.22 to 14.77, with five out of the six measures above 10.00. The results are summarized in Table IX.

TABLE IX

Showing the Per Cent of Reduction of the Overlapping in the Total Composite as Compared with the Average of 22 Tests

Comparisons	Average of 22 tests	Total composite	Per cent of reduction
VI on VIII	19.35	4.58	76
VI on VII	31.59	19.57	38
VII on VIII	34.96	27.74	21
VIII on VI	20.82	8.37	60
VIII on VII	34.33	24.15	30
VII on VI	33.08	21.98	34

Table IX reveals the amount of error that would have resulted had we used the results from single tests as a measure of overlapping of Grades VI, VII, and VIII. The error would have been greatest in the VI-on-VIII comparison and least on the VII-on-VIII, as shown by the fact that the per cent of overlapping is reduced 76 per cent in the former, and 21 per cent in the latter. The facts of Table IX show very clearly the limitations of the results from single tests as measures of overlapping, and illustrate the necessity of careful scrutiny of any general statements regarding the amount of overlapping of grades based upon such data.

Our results permit an answer to the question, How do the different grades tested compare with each other as to amount of overlapping? As might be expected, the overlapping in the case of the sixth and eighth grades is least. It might not have been anticipated, however, that the overlapping in the case of the seventh and eighth grades would exceed that of the sixth and seventh. Nevertheless it does so generally throughout.

Inspection of Table VII reveals the fact that the overlapping of Grades VII and VIII exceeds that of Grades VI and VII in 12 out of the 20 cases, upward and downward. This excess in the per cent of overlapping of Grades VII and VIII is brought

out more markedly when the magnitude of these differences is taken into account. The totals show an excess for Grades VII and VIII over Grades VI and VII of 96.45 as against 38.15 excess for Grades VI and VII over VII and VIII. That is, on the basis of the results by composites, Grades VII and VIII show an amount of overlapping in excess of that of Grades VI and VII about three times as great as that of Grades VI and VII over VII and VIII. A similar relation, though in reduced amount, holds in the results from single tests, as may be seen from Table II.

These facts seem to furnish some substantial evidence in support of the generally accepted notion that the work of the seventh and eighth grades is less clearly differentiated than that of the sixth and seventh. They appear to support the contention that the eighth grade is largely only a transition period. If so, they add weight to the argument for modification of our school grading scheme. It may also prove to be significant, especially for departmental teaching, that although Grades VII and VIII show more overlapping on the whole, Grades VI and VII show an excess in Composition, Trabue and Mixed Relations. (See Table VII.)

Summarizing, we may say that the amount of overlapping measured by a single test is reduced appreciably when measured by a composite of a number of tests of the same trait, there being a direct relation between the number of tests entering into the composite and the amount of reduction. Moreover, when a further grouping of scores is made by recombining these composites of single tests in a given trait into a gross composite representative of ability in a more complex trait, as, for example, control of the vernacular as exemplified in our English composite of 22 tests, the amount of overlapping is appreciably reduced. Extending this grouping so as to include a composite of results from tests of a still different trait, such as arithmetic, we get still more reduction of the overlapping in this complex, more nearly representative of school ability.

2. Surfaces of Overlapping

Figs. 5 to 10 show the distribution surfaces and resulting overlapping of Grades VI, VII, and VIII in several of the composites, as an aid in understanding the facts of Table VII. The arrangement of the surfaces is similar to that of Figs. 1 to 4, previously

explained. The figures are in order from greatest to least overlapping, roughly shown by the differences between the medians. Fig. 11 has been drawn in the form of smooth curves as a basis for comparison. The lower set of curves shows the amount of overlapping of Grade VI on Grade VIII in the total composite, 4.58 per cent, in contrast with the upper set of curves which shows the amount of overlapping of these grades in the average for 22 tests, 19.35 per cent. (See Table VIII.) The figure thus illustrates the amount of reduction by the composite method over the results from single tests.

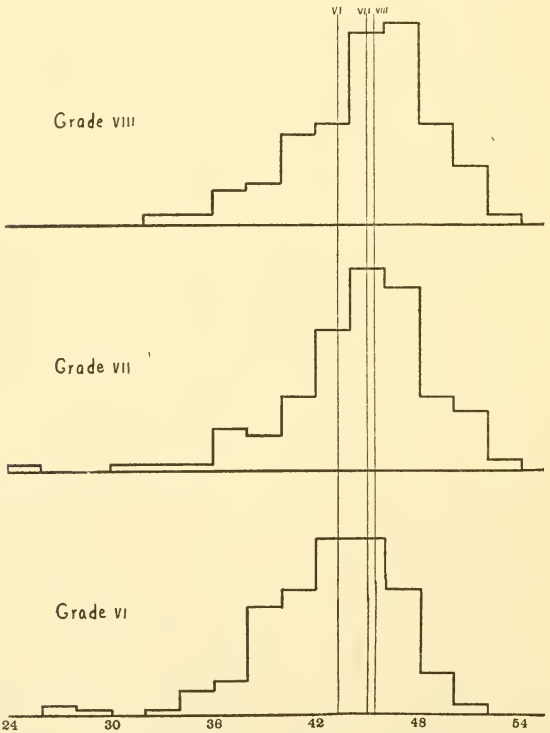

Fig. 5. Overlapping in Directions Composite.

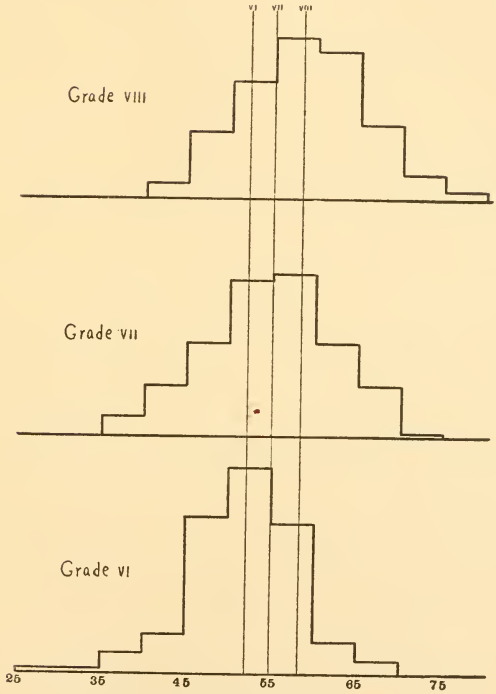

Fig. 6. Overlapping in Trabue Composite.

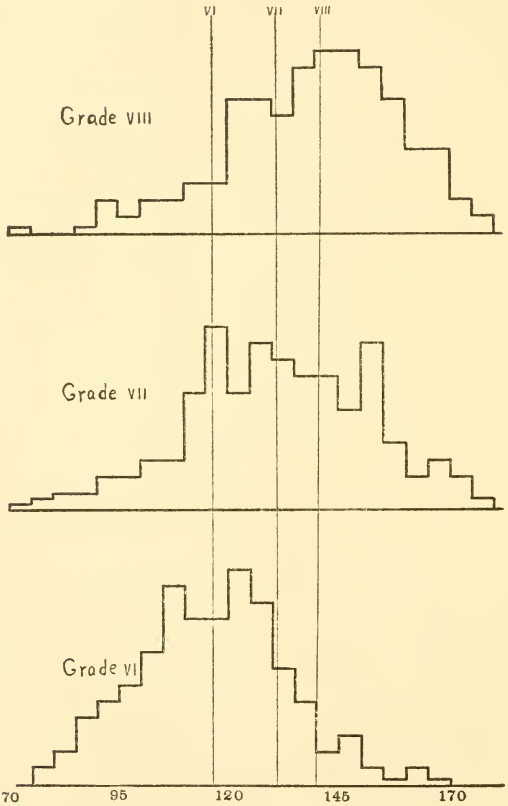

Fig. 7. Overlapping in Arithmetic Composite.

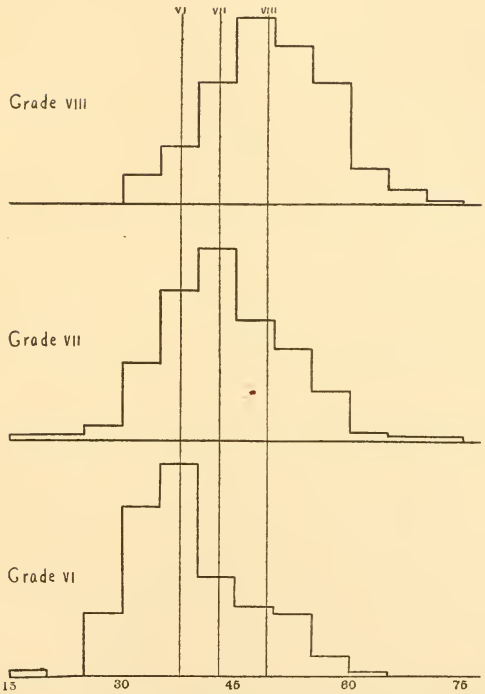

Fig. 8. Overlapping in Composition.

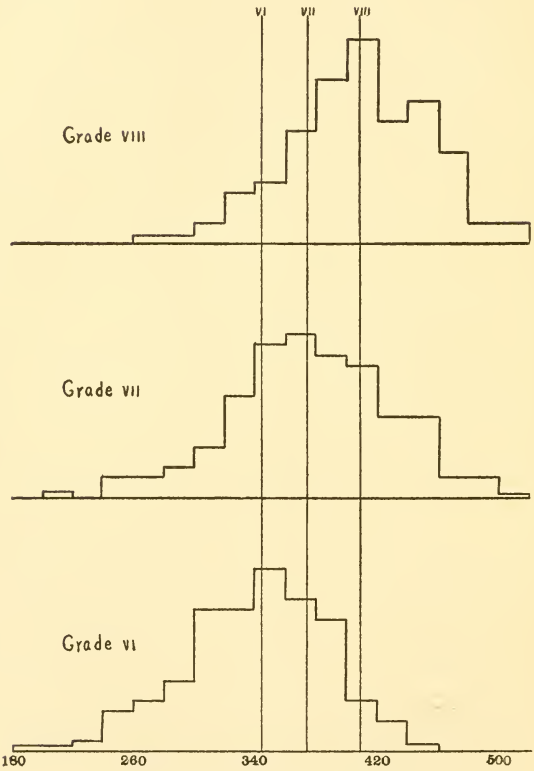

Fig. 9. Overlapping in English Composite.

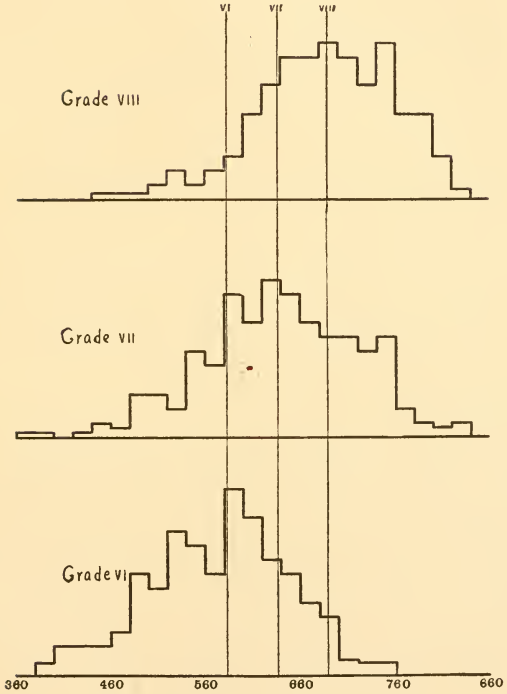

Fig. 10. Overlapping in Total Composite.

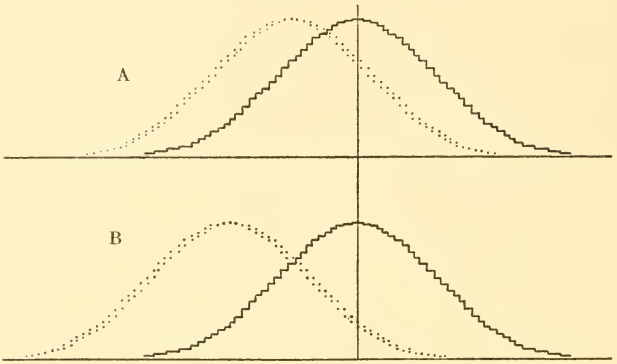

Fig. 11.

A. Overlapping of Grade VI on Grade VIII in the Average for 22 Tests, 19.35 Per Cent. (See Table VIII.)

B. Overlapping of Grade VI on Grade VIII in the Total Composite, 4 58 Per Cent. (See Table VIII.)

3. Statistical Treatment

Weighting of Tests.

In making the first group of composites each test entering into a given composite was counted equal in value to every other. For example, Addition, Subtraction, Multiplication, Division, Problems I, and Problems II were counted of equal value as units in the arithmetic composite. Therefore, the scores of a given pupil in these six tests were added and the sum taken as his score in arithmetic. Similarly in Trabue the scores in the four tests were added, and in Opposites the scores for three tests. A B–C composite was made by adding the scores in B1, B2, C1, C2, C3. Mixed Relations is a composite of D1 and D2; Directions of X, VI, and VII; and Vocabulary of VIII and XIII.

However, in making the composite for all the English tests it became necessary to recognize statistically what common sense demands, that one of these composites may have more or less value than another as a unit in a total English composite. Likewise in the final total, including all tests of both groups, Arithmetic and English, each composite had to be given a value.

This involved three steps: (1) calculating a measure of variability for each group of tests and each single test in each grade; (2) determining upon the relative weight to be given to each composite group or single test in a composite; (3) multiplying or dividing (as the case required) the scores in a given composite by the factor necessary to bring them to a basis comparable with all the others.

It was found that the standard deviations (the measure of variability used) for the three grades varied one with another throughout the different tests so little that the average of the three S. D.'s for a given test was used.

In determining upon the weights to be given to the different tests the combined judgment of eight judges, familiar with the tests, was used. Table X shows the weights given and the factor by which the scores in each group of tests or in each test were multiplied or divided.

TABLE X

WEIGHTS GIVEN TO SINGLE TESTS AND TO GROUPS OF TESTS ENTERING INTO THE ENGLISH COMPOSITE AND THE TOTAL COMPOSITE

	No. of tests	S. D. Average for the 3 grades	Multiple			Weight given
Spelling	1	4.43	×	1	=	4.43
Alpha 2	1	3.84	×	2	=	7.68
Composition	2	8.07	×	1	=	8.07
Trabue	4	7.17	×	2	=	14.34
Opposites	3	54.13	÷	5	=	10.83
B–C	5	17.61	÷	4	=	4.40
Mixed Relations	2	16.51	÷	2	=	8.26
Directions	2	4.09	×	1	=	4.09
Vocabulary	2	39.70	÷	5	=	7.94
Arithmetic	6	19.20	×	2	=	38.40

Interpolated Scores.

The following rules were followed in supplying missing scores:

1. In arithmetic interpolations were made if not more than three out of the six scores were lacking. The median score for the particular class was used, rather than the median for the whole grade.

2. If only one composition was written the score for this was neglected and the median for the class for both compositions was used.

3. In the B–C composite interpolations were made when not more than two scores out of the five were missing.

4. In the Mixed Relations composite when one score was missing the score for the other test was used, as the central tendencies for the two tests are practically equal.

5. In all cases of single tests, e.g., Alpha 2 and Spelling, the medians for each class were used.

6. In the total English composite interpolations were made if not more than four out of the nine scores were missing.

7. In the total composite a missing score in mathematics was interpolated when not more than two others, in the English group, were interpolated; five scores were interpolated in the English group if the mathematics score was present.

V

RELIABILITY OF THE TESTS USED AS DETERMINED BY THE CORRELATION OF GROUPS OF SIMILAR TESTS

So far we have been attempting to measure the amount of overlapping of attainments in certain sixth, seventh, and eighth grades by the use of more or less well standardized tests. It is our purpose in this chapter to apply a measure of reliability to the tests we have used, the results of which should serve the twofold purpose of measuring the validity of our results and serve as a guide to any who may wish to know how much reliance may be placed in certain tests as measures of attainment in Grades VI, VII, VIII.

Most of the work heretofore done in the way of measuring the reliability of given tests as measures of general mental ability has been by means of correlations among the tests, but with single or, at most, very few measures by a given test. Our data furnish a particularly good opportunity to supplement these results, as we have the advantage of comparison by composites, thus tending to overcome the error resulting from the use of only one or two tests of a given trait. Further we have the very real advantage of large numbers. While this latter gives greater validity to the results obtained, it makes the calculation of many coefficients of correlation prohibitive because of the amount of labor involved. We have chosen one comparison of two groups of tests in arithmetic and one of two groups of tests in English.

Two Arithmetic composites were made by adding the scores in addition, division and the first problem test for the one, and subtraction, multiplication, and the second problem test for the other. Likewise two English composites were made. The first was composed of the scores in the Trabue, Opposites, B–C, and Visual Vocabulary composites and the second, of those in the

Composition, Reading, Mixed Relations, and Directions composites. Each was weighted as indicated above.

Coefficients of correlation were then calculated by the Pearson method, formula $r = \frac{\Sigma(x.y)}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}}$, for each group of composites, for each grade, with the results indicated below.

TABLE XI

COEFFICIENTS OF CORRELATION BETWEEN ARITHMETIC COMPOSITES AND BETWEEN ENGLISH COMPOSITES

	ARITHMETIC		ENGLISH	
	One group of three tests with another group of three		One group of five tests with another group of five	
Grade	No. of pupils	r	No. of pupils	r
VI	240	.706	256	.739
VII	310	.774	325	.749
VIII	275	.713	283	.705
Average		.731		.731

From these coefficients of correlation we are able, by the use of Brown's formula (3), to get a measure of the reliability of the Arithmetic composite and of the English composite as measures of these abilities in the grades studied. Applying the formula, $r_n = \frac{n(r_1)}{1 + (n - 1)r_1}$, we get the following results:

TABLE XII

Grade	Approximate coefficient of reliability of the Arithmetic composite as representative of all the Arithmetic work of these grades	Approximate coefficient of reliability of the English composite as representative of all the English work of these grades
VI	.828	.850
VII	.873	.856
VIII	.832	.827
Average	.844	.844

That is, on the basis of the correlation between the two Arithmetic composites the coefficient of reliability of the whole Arithmetic composite as a measure of the arithmetic work of these grades is .844, as an average for the three grades. It happens that the coefficient of reliability of the English composite is exactly the same.

VI

COMPARISON OF RESULTS IN OVERLAPPING WITH THOSE OF OTHER INVESTIGATORS

So far as the writer is aware, no previous investigation has had as its special aim the study of overlapping. Many students of individual differences have reported overlapping figures on various bases, such as grade in school, age, sex, and race. Practically all the studies of achievements of school pupils, for whatever purpose, call attention to the overlapping of grades as a significant finding. It is the purpose of this chapter to report some of these results. In all cases where the data permit, per cents of overlapping have been calculated on the basis used in this study, where this had not already been done. Results for the sixth, seventh, and eighth grades only, where these grades are included in the data available, will be given, the purpose being to make as direct a comparison as possible. Some quotations will be made from these studies to show the nature of the conclusions based upon the extent of the overlapping.

Chambers, in a study of "Individual Differences in Grammar Grade Children" (7), made in 1910, treats the question of overlapping of grades at considerable length. He gave one test each in cancellation of A's, addition, spelling, association of opposites, auditory memory, and visual memory to 22 pupils in a seventh grade and the same number in an eighth grade. Below are the per cents of overlapping calculated from his data.

Per cents of overlapping	Ca	Ad	Sp	As	Am	Vm
of Grade VII on Grade VIII....	47	43	55	55	41	52

The author says on page 69, "Manifestly an extra year of school life has failed to produce any discernible improvement in the

traits tested. Indeed in two of the abilities concerning which the public is most insistent as to the school's responsibility, viz., addition and spelling, the extra year in school seems to have produced a positive deterioration: in addition the lowest rank is monopolized by a single eighth grade representative while the two highest ranks are preempted by two members of the seventh grade. In spelling the lowest rank is occupied by four eighth and one seventh grade pupils, while the highest rank contains one from each grade." Further, on page 71: "The most important conclusion reached in the comparison of the distribution of abilities in the two grades is, then, that there is no line of demarcation between them; in regard to every trait examined the grades overlap and in regard to most traits they are coextensive in their range. Hence, to assume that in two school grades we have two distinct species, that certain abilities are lacking in one and present in the other, that all the members of one class are of approximately equal ability in a certain field and that they are all inferior in that ability to all the members of the other, is, at best, a very hazardous guess . . . Teachers cannot afford to forget that our school grades do not represent distinct gradations of ability in the pupils, but are simply convenient devices of administration to facilitate the handling of children in the mass. The grades are determined by reference to more or less artificial standards, and too often do not represent the real intelligence, industry, endurance, adaptability, and other traits important for education of the pupils. It seems as if ability in perception, association and memory, when these functions are tested on familiar material, should be pretty closely related to educability, and should be affected by school progress to an extent distinguishable in successive grades, if the grades really grade."

Bonser (2) presents overlapping figures on the basis of both ages and grades, for each of seven tests of reasoning ability of children in the fourth, fifth and sixth grades, and also for the combined results of the seven tests. Per cents of overlapping calculated from his data are given in Table XIII.

TABLE XIII

PER CENT OF BOYS IN ONE GRADE WHO EQUAL OR EXCEED THE MEDIAN OF THE OTHER GRADE

Tests	VB on VIB	VA on VIA
I and II—Arithmetic	25.46	32.31
III—Controlled Association	35.95	30.43
IV—Opposites	30.38	20.00
V—Selective Judgment	33.23	19.23
VI—Selective Judgment	21.84	30.77
VII—Intellectual Interpretation of Poems	41.71	24.46
Average	31.43	26.20
Combined Results	23.54	7.69

We have further evidence here of the high per cents of overlapping by single tests and of the reduction of the overlapping by a composite of all the scores from that obtained by averaging the per cents for the different tests. Regarding the method of obtaining the composite the author says on p. 72, "The sums of the results for the several tests have been taken and arrayed by the same distributions as the results of each separate test."

To compare further the overlapping figures by single tests and by composites Table XIV has been prepared by assembling the results for the different tests and for the composite as calculated by Bonser.

TABLE XIV

PER CENT OF BOYS IN EACH GRADE WHO REACH OR EXCEED THE ABILITY REACHED BY THE HIGHEST 25 PER CENT OF ALL THE BOYS TOGETHER

Tests	IVA	VB	VA	VIB	VIA
I and II	5.37	12.65	32.30	44.89	62.00
III	17.20	22.77	27.68	42.85	52.00
IV	7.52	18.97	24.56	42.85	54.00
V	7.52	29.09	16.91	39.79	42.00
VI	15.05	22.77	24.56	54.08	48.00
VII	16.12	20.24	16.91	37.75	56.00
Average	11.46	21.08	23.82	43.70	52.33
Combined Results	9.67	18.97	19.95	42.85	62.00

Here again we note a reduction of the figures for the lower grades in favor of the composites. It will be observed that in

this table the higher figure for VIA, 62.00 for the composite as compared with the average, 52.33, argues in favor of the composite also, because the larger the per cent in the sixth grade the more superior is the sixth-grade attainment to that of the lower grades and hence the less the overlapping.

Bonser gives the following as one of his "General Conclusions": "The point of greatest pragmatic significance for the school lies in the implications from the two facts, first, that there are quite substantial percentages from both the lower grade groups and lower age groups who are found in the highest quartile of ability for all; second, that most of the groups of the youngest 25 per cent in each grade show higher ability than the oldest 25 per cent and sometimes higher than that of the median ability of the whole grade." (p. 91)

Thorndike (18) in reporting Bonser's study makes the following comments. "It should be borne in mind, however, that (except with the 'opposite' test) the time allowed in each grade was not necessarily identical, each class being given such time as the quickest person in it required to complete the test. Bonser does not regard the time factor as of much consequence, in view of the nature of the tests, but it seems probable that the lower grades had longer time and so are credited with somewhat better relative scores than they would have obtained if all grades had been given in every test some constant time." (p. 64) After making these qualifications, Thorndike gives some striking interpretations of the overlapping figures. "If we drew at random 109 boys and girls from the 757 in all these grades to make up the VIA, *this absolutely random drawing* would differ from the IVA grade by half as much as does the group picked out administratively as two years in advance of it . . . Indeed I unhesitatingly assert that a month's test in respect to the ability to do the specific intellectual work of the school course of study would show a similar, though perhaps not so great, variability and a similar overlapping." (p. 66)

Courtis (8) presents figures which show that 2.8 per cent of the fourth grade reach or exceed the median of the eighth, and 4.1 per cent of the eighth go below the median of the fourth in the Courtis arithmetic tests.

In his report of the use of the Courtis tests in the New York Inquiry (9) Courtis gives considerable space to the question of

overlapping. Typical results are as indicated below, calculated from data on page 441.

28.9 per cent of the 6th grade reach or exceed the median of the 8th.
40.6 per cent of the 6th grade reach or exceed the median of the 7th.
36.1 per cent of the 7th grade reach or exceed the median of the 8th.

Courtis makes the following comments: "So far as any individual child is concerned, to say that he has completed the course in arithmetic in the public schools is to convey *no* information as to his ability in even the simplest work." (p. 439) "The generalization to be made from this is that the amount of overlapping of the grades is constant, and is, therefore, due to the one factor that is common to all the schools and grades—that is, to the inherent differences in children in their ability to respond to training in multiplication tables." (p. 450)

The results obtained by Elliott (10) from single tests vary considerably among schools. He found that the per cents of overlapping of all fifth grades on all seventh grades were as given below.

School Systems	Arithmetic	Composition	Spelling
A, B, C	18.80	4.50	19.90
D	13.60	17.30	34.60
E	25.00	0.00	27.42
F	2.56	4.87	11.25
G	26.00	55.10	27.45

Assembling Starch's figures as reported in Part I of the Fifteenth Yearbook of the National Society for the Study of Education (17), we get the following.

Test	Per Cent in any Grade Who Equal or Exceed the Median of the Next Grade Above
Reading	31
Spelling	23
Arithmetic	32.5

With regard to a composite this author says, "A combined score for all studies was computed for each pupil so that the various subjects were balanced against each other. It was found that even then the overlapping was practically as large, Thirty-two and two-tenths per cent of the pupils in any given grade reach or exceed the standard of the next grade above it; thirty-five and two-tenths per cent fall to or below the standard of the next

grade below." (p. 145) In the absence of distribution tables and any further statement of the method employed it is not possible to make any careful comparison with our figures.

From a five-minute reading test, scored on the basis of reproduction of matter read and answers to ten questions, Waldo (26) found that "many children in lower grades excel their schoolmates in higher grades. Thus 29.5 per cent of the fifth graders excel the average rate in reading of the eighth grade; 36.1 per cent of them excel the seventh grade average, and 42.6 per cent excel the average for the sixth grade.

"In reproduction there are smaller numbers of children who are superior to those of higher grades; but 16 children, or 26.2 per cent, excel the average of the sixth grade; and but 8, or 15.7 per cent, exceed the seventh grade average, and none are superior to the eighth."

Thorndike (22) found that in handwriting "individual pupils within the same grade . . . show a range of difference much greater than that between the fifth grade of the worst system and the eighth grade of the best."

Buckingham (5) reports results of tests given in New York City in 1915, as follows:

PER CENT IN EACH GRADE WHO EQUAL OR EXCEED THE MEDIAN OF THE NEXT HIGHER GRADE

Tests	VII[1] on VII[2]	VII[2] on VIII[1]	VIII[1] on VIII[2]
Arithmetic	33	36	28
Geography	47	39	28
History	43	23	27
Grammar	?	38	29

Buckingham gave a test consisting of ten problems in arithmetic to 4985 children in March and again a similar test in June, from the results of which he got per cents of overlapping as shown below (4).

	VII[1] on VII[2]	VII[1] on VIII[1]	VII[1] on VIII[2]	VII[2] on VIII[1]	VII[2] on VIII[2]	VIII[1] on VIII[2]
March	37.3	25.3	11.3	36.2	19.9	30.0
June	32.2	28.6	13.7	45.5	22.6	28.5

Table XV brings together the results of a number of other studies.

TABLE XV

SUMMARY OF RESULTS OF OTHER INVESTIGATIONS

PER CENT IN EACH GRADE WHO EQUAL OR EXCEED THE MEDIAN OF THE OTHER GRADE

Investigators	Tests	No. of pupils	VI on VIII	VI on VII	VII on VIII
Buckingham (6)	Spelling 100 words	1060	13.6	34.0	27.1
Buckingham (6)	Spelling 100 words	1940	16.8	36.1	26.8
Gray (11)	Quality of Silent Reading	1053	28.8	38.5	38.4
Trabue (25)	Completion Test Scale A	4337	12.2	25.4	31.3
Studebaker (19)	Courtis Series B				
	Addition	2936	31.4	41.3	38.7
	Subtraction	2930	25.3	40.8	34.8
	Multiplication	2941	25.2	42.8	35.1
	Division	2941	21.2	36.5	30.3
Salt Lake Survey (15)	Stone Reasoning	1161	7.7	25.7	34.8
	Composition	1667	20.6	38.9	32.4
	Penmanship	1736	23.2	30.2	39.2
	Speed of Silent Reading	1165	52.0	43.7	59.2
Kelly (13)	Kansas Silent Reading 2nd test	3514	27.4	37.0	42.2
Butte Survey (14)	Courtis Series B Multiplication	414	10.5	26.1	26.4
	Stone Reasoning	416	11.2	24.2	24.4

The one outstanding fact in all this is the great amount of overlapping of grades when the pupils are tested with single or few tests. In the light of our discussion in Chapter IV, however, these large

overlapping figures are not so disturbing as they might otherwise be. That many pupils in one grade excel the median pupil of a higher grade in a given test is not greatly significant, really telling us very little about the extent to which the grades overlap in ability in a given trait.

Overlapping of Attainments on the Age Basis

While it is the purpose of this study to consider only the problem of overlapping of attainments in certain grades, it should be pointed out that this represents only one basis for the study of overlapping. Another very important question is the overlapping of the attainments of children of different age groups. Our data have not been treated in such a way as to show what the results from our testing would be. Table XVI has been prepared by bringing together figures in Bonser's study, in order to show the facts as he found them. The table is similar to Table XIV with the exception that here our basis is age rather than grade.

TABLE XVI

PER CENT OF BOYS IN EACH AGE GROUP WHO REACH THE ABILITY REACHED BY THE HIGHEST 25 PER CENT OF ALL THE BOYS TOGETHER

Tests	Age Groups 8–11	11–12	12–13	13–16
I and II	27.02	26.13	34.37	29.87
III	27.92	39.60	22.90	35.04
IV	29.72	32.67	27.06	19.47
V	28.82	31.68	28.10	20.76
VI	25.22	35.64	29.14	22.06
VII	27.02	27.72	29.14	25.96
Average	27.62	32.24	28.45	25.53
Combined Results	26.12	29.70	21.86	23.36

It is seen that roughly one-fourth of the boys in each of these four age groups reach the 75 percentile for the whole group. That is, the highest 25 per cent of the whole group is composed of about equal parts of each age group. Further it may be noted that the combined results are consistently lower than the averages of the single tests, and in one case, the 12–13 group, considerably lower.

VII

CONCLUSIONS

1. This study supports the findings of other investigators in the generally high per cents of overlapping among the upper elementary school grades when measured in terms of attainments in single tests.

2. Recognising the limitations of such testing and letting the figures stand for just what they are, per cents of overlapping in the tests as given, we find that those tests which show the least amount of overlapping are, in order, Spelling, Problems I, Visual Vocabulary VIII, Visual Vocabulary XIII, Opposites A4, Opposites A2, and Alpha 2.

3. The large per cents of overlapping from single tests are appreciably reduced when the overlapping is based on the results from composites of two or more tests of a similar kind.

4. There is still further reduction of the per cents of overlapping when these composites are further combined, so as to be more nearly representative of general ability to do the work of a given grade.

5. It follows that the error in using overlapping figures from single tests as measures of the overlapping in a given trait is great. It is, of course, still greater if such figures are used as measures of general ability to do the work of the grade, even though a test be given in each study.

6. Our grading system as a means of placing children according to ability to do the work of the school is not so grossly lacking as has been suggested on the basis of overlapping by single tests. On the other hand, our school grades do not represent distinct types of ability, so clearly marked off from each other that all the pupils of a given grade are superior to all the pupils of the next lower grade and inferior to all the pupils of the next higher grade in ability to do the work of the school.

7. The fact that the overlapping of Grades VII and VIII is in general higher than that of Grades VI and VII is further evidence in support of the belief that the eighth grade as at present maintained in the ordinary eight grade scheme does not justify itself.

8. Statements regarding the futility of a given half-year or year of school work are not sufficiently well founded on the basis of the large amount of overlapping of grades in certain tests.

9. The arithmetic tests used in this study, treated by the composite method, have a high degree of reliability as a measure of the arithmetic work of these grades. The same is true of the English composite, the coefficient of reliability being .844 in each case, as an average for the three grades.

VIII

BIBLIOGRAPHY

1. Ayres, Leonard P. Measuring Scale for Ability in Spelling. Russell Sage Foundation, New York City, Division of Education.

2. Bonser, Frederick G. The Reasoning Ability of Children of the Fourth, Fifth, and Sixth School Grades. Teachers College, Columbia University, Contributions to Education, No. 37.

3. Brown, William. Essentials of Mental Measurement, p. 102.

4. Buckingham, B. R. Notes on the Derivation of Scales in School Subjects. Fifteenth Yearbook of the National Society for the Study of Education, 1915, Part 1, p. 33.

5. Buckingham, B. R. Principles of Scale Derivation with Special Application to Arithmetic, Geography, History and Grammar. Third Annual Conference on Educational Measurements Bloomington, Indiana, 1916, p. 83.

6. Buckingham, B. R. Spelling Ability: Its Measurement and Distribution. Teachers College, Columbia University, Contributions to Education, No. 59, pp. 32, 68.

7. Chambers, Will Grant. Individual Differences in Grammar Grade Children. *Journal of Educational Psychology*, Vol. 1, 1910, pp. 61–75.

8. Courtis, S. A. Educational Diagnosis. *Educational Administration and Supervision*, Vol. 1, 1915, p. 93.

9. Courtis, S. A. The Courtis Tests in Arithmetic. New York School Inquiry, Vol. I, pp. 391–546.

10. Elliott, C. H. Variations in the Achievements of Pupils. Teachers College, Columbia University, Contributions to Education, No. 72, p. 51.

11. Gray, William Scott. Studies of Elementary School Reading through Standardized Tests. Supplementary Educational Monographs, No. 1, University of Chicago Press, p. 116.

12. Hollingworth, H. L. Correlation of Abilities as Affected by Practice. *Journal of Educational Psychology*, Vol. 4, 1913, p. 411.

13. Kelly, F. J. The Kansas Silent Reading Tests. *Journal of Educational Psychology*, Vol. 7, 1916, p. 77.

14. Report of A Survey of the School System of Butte, Montana, 1914.

15. Report of a Survey of the School System of Salt Lake City, Utah, 1915.

16. Starch, Daniel. Educational Measurements.

17. Starch, Daniel. Standard Tests as Aids in the Classification and Promotion of Pupils. Fifteenth Yearbook of the National Society for the Study of Education, 1915, Part 1.

18. Strayer and Thorndike. Educational Administration.

19. Studebaker, J. W. Annual Report of the Des Moines Public Schools, 1915, pp. 62–5.

20. Thorndike, E. L. An Improved Scale for Measuring Ability in Reading. *Teachers College Record*, Vol. XVI: No. 5, Nov. 1915.

21. Thorndike, E. L. Educational Psychology, Vol. III, pp. 177, 180–181.

22. Thorndike, E. L. Handwriting. *Teachers College Record*, Vol. XI: No. 2, March 1910, p. 75.

23. Thorndike, E. L. Preliminary Extension of the Hillegas Scale for the Measurement of Quality in English Composition by Young People. Teachers College, Columbia University.

24. Thorndike, E. L. The Measurement of Ability in Reading. *Teachers College Record*, Vol. XV: No. 4, Sept. 1914.

25. Trabue, M. R. Completion-Test Language Scales. Teachers College, Columbia University, Contributions to Education, No. 77.

26. Waldo, Karl D. Tests in Reading in Sycamore Schools. *Elementary School Journal*, Vol. 15, 1914–15, p. 261.

27. Woody, Clifford. Measurements of Some Achievements in Arithmetic. Teachers College, Columbia University, Contributions to Education, No. 80.

IX

APPENDIX—DISTRIBUTIONS

Addition

Score	Grade VI	Grade VII	Grade VIII
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15	1	1	
16		1	
17	1		1
18		3	
19	2	4	
20	2	4	
21	7	3	3
22	4	7	6
23	9	7	3
24		10	10
25	15	2	13
26	12	23	2
27	27	24	10
28	23	18	1
29	27	29	23
30	16	20	21
31	19	33	26
32	20	17	27
33	16	29	26
34	15	14	27
35	4	22	27
36	5	17	22
37	2	71	15
38	2	4	2
No. tested	236	306	273
Median Score	29.30	30.89	32.35

Subtraction

Score	Grade VI	Grade VII	Grade VIII
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15		1	
16			
17	1	2	
18	1	1	1
19	7	1	5
20	6	2	1
21	13	16	6
22	17	14	7
23	23	12	13
24	28	17	8
25	31	16	18
26	24	25	19
27	20	31	21
28	16	35	43
29	24	36	31
10	9	21	39
32	8	27	23
33	6	60	14
33	3	14	90
34	1	15	17
35			31
No. tested	238	302	267
Median Score	25.74	28.37	28.82

DISTRIBUTIONS

Score	MULTIPLICATION Grade VI	VII	VIII	Score	DIVISION Grade VI	VII	VIII
0				0			
1				1			
2				2			
3				3			
4				4			
5				5			
6				6			
7				7			
8				8			
9				9			
10				10		1	
11				11			
12				12	1		
13				13			1
14		1		14	3	2	2
15			1	15	6	2	
16		1		16	3	5	
17	3	4	2	17	4	2	1
18	5	2	1	18	14	6	7
19	4	6	1	19	11	7	8
20	3	1	1	20	15	10	3
21	7	4	5	21	13	19	12
22	11	10	6	22	14	21	11
23	8	7	4	23	18	21	17
24	21	10	10	24	28	35	20
25	11	20	17	25	20	21	21
26	27	17	11	26	21	25	39
27	22	28	16	27	24	32	32
28	20	22	17	28	16	25	24
29	25	30	23	29	8	19	23
30	21	25	31	30	10	13	19
31	19	27	22	31	2	15	13
32	16	20	25	32	1	13	11
33	3	22	37	33	2	6	4
34	1	21	16	34		3	3
35	7	12	14	35		2	1
36	3	7	6	36		2	
37		5	6		—	—	—
38		2	1	No. tested	234	307	272
39		1					
	—	—	—	Median Score	24.54	26.06	26.85
No. tested	237	305	273				
Median Score	27.84	29.65	30.69				

Distributions

Problems I

Score	Grade VI	Grade VII	Grade VIII
0	16	3	
1	1		
2	2		
3	34	15	8
4	2	1	
5	6	3	2
6	36	25	12
7	7	2	1
8	10	4	1
9	37	33	19
10	6	1	2
11	10	7	1
12	24	36	21
13	3	3	6
14	7	7	3
15	10	38	41
16	2	5	2
17	3	8	2
18	9	44	42
19	2	6	2
20	2	9	7
21	5	28	41
22		3	3
23		1	
24	2	16	31
25		1	2
26			9
27		7	13
28			
29			
30		1	
No. tested	236	307	271
Median Score	9.11	15.36	18.35

Problems II

Score	Grade VI	Grade VII	Grade VIII
0	60	40	23
1	51	29	10
2	30	29	19
3	43	44	37
4	14	10	12
5	8	9	9
6	14	63	49
7	3	4	8
8		3	7
9	11	44	46
10	1	1	2
11	1		2
12	3	24	32
13			
14			1
15		4	13
16			
17			
18		1	3
No. tested	239	305	273
Median Score	2.28	5.06	6.54

Problems III

Score	VI	VII	VIII
0– 2	159	107	41
3– 5	40	66	52
6– 8	30	73	59
9–11	8	24	50
12–14	2	28	39
15–17	1	10	28
18–20		2	4
No. tested	240	310	273
Median Score	...	5.18	8.21

DISTRIBUTIONS

	TRABUE–B			TRABUE–C		
	Grade			Grade		
Score	VI	VII	VIII	VI	VII	VIII
0						
1						
2						
3						
4	1			1		
5						
6	1	1		1		
7		3	1	4	3	1
8	4	6	2	5	5	
9	5	5	2	12	14	4
10	23	17	7	9	24	10
11	19	27	14	40	34	23
12	50	38	38	53	44	33
13	51	49	46	44	48	35
14	44	39	44	34	40	47
15	16	47	38	17	38	34
16	12	44	28	18	27	40
17	14	18	26	2	10	18
18	4	5	16	4	15	17
19		2	1		1	8
20		1	7			3
No. tested	244	302	270	244	303	273
Median Score	13.37	14.13	14.57	12.94	13.57	14.65

DISTRIBUTIONS

	TRABUE-D			TRABUE-E		
	Grade			Grade		
Score	VI	VII	VIII	VI	VII	VIII
0						
1						
2						
3						
4						
5						
6						
7	2				1	
8	1	6	3	7	2	1
9	6	3	3	4	9	7
10	10	9	4	18	16	4
11	15	14	9	27	22	10
12	43	35	28	50	44	22
13	32	45	27	51	62	31
14	68	64	49	38	49	39
15	34	45	39	28	39	42
16	27	44	49	15	32	47
17	4	25	36	3	11	21
18	3	9	15	2	9	31
19		3	6		2	8
20			3		5	7
No. tested	245	302	271	243	303	270
Median Score	14.20	14.61	15.32	13.30	13.93	15.50

DISTRIBUTIONS

READING–ALPHA 2

Score	Grade VI	Grade VII	Grade VIII
0			
1			
2			
3			
4			
5			
6			
7			
8			
9	1		
10			
11	1	1	
12	2		1
13	4	1	
14	2	3	
15	5	3	
16	8	1	3
17	8	7	
18	14	10	3
19	23	9	6
20	18	14	5
21	27	21	5
22	18	19	19
23	22	26	14
24	17	27	26
25	23	33	21
26	25	32	30
27	15	31	38
28	6	34	43
29	7	11	27
30	1	12	15
31	1	11	13
32		1	3
33		1	
34			
35			
36			
37			
38			
No. tested	248	308	272
Median Score	22.61	25.37	27.08

DISTRIBUTIONS

Score	Opposites–A1 Grade VI	VII	VIII	Opposites–A2 Grade VI	VII	VIII
0	2	2	1	3	4	
5		1	1	1		
10			1	1	1	
15	1				1	
20				11	4	1
25				4	4	
30				15	7	
35		2		16	9	3
40				19	9	5
45	1			33	18	14
50				29	22	21
55	1	1		32	28	23
60	1		1	31	42	25
65				26	27	23
70		1		9	30	32
75	3			8	34	27
80	1			3	24	22
85	2	1		5	22	32
90		1	3	1	11	15
95	5	5	1	2	6	14
100	5	5	1		2	11
105	10	5	2		4	1
110	13	11	5		1	1
115	36	26	18			2
120	80	72	43			1
125	61	84	81			1
130	20	67	77			
135	6	22	41			
No. tested	248	306	276	249	310	274
Median Score	122.69	126.19	128.77	53.71	66.11	73.44

DISTRIBUTIONS

	OPPOSITES–A3			OPPOSITES–A4		
	Grade			Grade		
Score	VI	VII	VIII	VI	VII	VIII
0	7	4		5	4	
5	2	2	1	3	1	
10	2	2		6	4	2
15	7	1		10	3	
20	9	4	2	10	4	2
25	9	6		17	9	2
30	17	11	7	19	7	6
35	25	10	9	20	15	2
40	28	26	7	21	21	10
45	21	37	14	21	27	9
50	36	38	24	17	27	14
55	25	35	27	27	20	22
60	17	31	32	19	15	23
65	19	16	22	20	33	28
70	8	16	20	16	37	23
75	6	19	30	9	25	25
80	6	16	21	6	14	29
85	3	9	22	1	22	19
90	1	5	11	1	4	15
95	1	3	13	2	7	14
100	1	10	9		8	13
105		2	1	1	6	5
110			1		1	5
115					1	2
120						1
125						
130						
135						
No. tested	250	303	273	251	315	271
Median Score	49.52	56.50	68.07	48.45	65.08	73.37

DISTRIBUTIONS

	WHOLE–PART–B1			ADJECTIVE–SUBSTANTIVE–B2		
	Grade			Grade		
Score	VI	VII	VIII	VI	VII	VIII
0	3	1	6	18	12	9
1	7	12	11	13	11	5
2	11	14	10	3	6	1
3	16	4	10	3	1	3
4	10	10	10	1	2	2
5	4	7	6	1	1	2
6	4	5	8	2	1	
7	4	7	7			
8	2	7	5	1		
9	8	5	7	1		2
10	10	4	3			1
11	6	4	4	2	1	
12	8	6	4	1	1	
13	2	7	4	2	1	1
14	9	5	10	4	1	2
15	15	11	6	2	2	1
16	15	13	17	3	1	2
17	18	26	15	9	10	11
18	22	31	33	11	14	11
19	41	56	47	38	48	38
20	38	79	50	134	201	182
No. tested	253	314	273	249	314	273
Median Score	16.50	18.29	17.57			

DISTRIBUTIONS

Score	VERB–OBJECT–C1 Grade VI	Grade VII	Grade VIII	SPECIES–GENUS–C2 Grade VI	Grade VII	Grade VIII
0	3	2	1	1		
1	1	1				1
2	1			1		
3	1			3		
4		1		1	3	
5				3	2	1
6				2	4	2
7				4		1
8				5	4	2
9	2	1		4	4	1
10				4	4	
11				3	2	2
12				3	4	2
13	1			3	5	3
14		1		2	3	4
15			1	3	3	5
16			1	2	1	1
17		3		1	4	4
18		2		4	5	
19	2	1		3	4	3
20	5	3		5	1	5
21		1		9	5	6
22	2	4	1	4	11	4
23	3	2	3	8	14	6
24	7	6	1	18	5	8
25	8	7	3	19	15	12
26	15	14	11	24	36	12
27	13	25	18	26	32	34
28	37	31	22	39	51	57
29	71	82	66	30	50	62
30	70	25	145	9	40	35
No. tested	242	312	273	243	312	273
Median Score	29.28	29.62		26.27	27.56	28.31

DISTRIBUTIONS

	SPELLING			PART-WHOLE-C3		
	Grade			Grade		
Score	VI	VII	VIII	VI	VII	VIII
0	1	1			1	
1	6			1	1	
2	4	3				
3	8	5			3	
4	7	4		2	9	3
5	10	4	1	4	1	2
6	11	6		6	3	1
7	11	6	1	2	5	2
8	20	2	4	8	8	5
9	13	8	1	5	5	9
10	17	11	3	15	10	10
11	10	10	5	11	17	4
12	11	10	7	19	21	13
13	11	12	6	17	22	11
14	19	21	7	20	24	31
15	16	20	12	29	35	22
16	21	24	20	24	51	39
17	19	33	31	29	34	46
18	19	33	39	26	26	34
19	8	44	61	15	27	27
20	10	47	74	7	7	18
No. tested	252	304	272	240	310	274
Median Score	12.73	17.15	18.97	15.34	15.71	16.69

DISTRIBUTIONS

	MIXED RELATIONS–D1			MIXED RELATIONS–D2		
	Grade			Grade		
Score	VI	VII	VIII	VI	VII	VIII
0	2	5	3	2	3	
2	3	2	2	1	4	1
4	6	1	2	1	2	4
6	7	11	6	11	9	6
8	11	20	13	12	21	5
10	14	24	20	23	39	11
12	31	43	33	34	36	18
14	48	56	38	26	26	34
16	51	29	46	26	22	26
18	22	30	22	11	17	22
20	14	19	9	6	9	8
22	11	10	5	6	6	12
24	7	7	7	4	14	6
26	5	7	10	7	13	11
28	4	7	5	10	5	7
30	6	14	10	20	10	8
32	3	6	13	20	16	18
34	3	11	14	19	19	19
36	2	5	7	8	30	34
38		2	8	2	10	22
40		2		2	1	1
No. tested	250	311	273	251	312	273
Median Score	16.12	15.76	16.85	17.19	17.45	22.25

DISTRIBUTIONS

	DIRECTIONS–X			DIRECTIONS–VII		
	Grade			Grade		
Score	VI	VII	VIII	VI	VII	VIII
0	1					
1	1			1		
2	1	1		1	3	6
3		1				
4		3		1		
5			1			
6	3	4		3	3	2
7	2	2	2	12	9	6
8	5	2	5	31	24	17
9	6	5	3	62	67	34
10	12	4	6	74	82	72
11	13	11	12	29	38	45
12	30	18	25	20	41	46
13	16	14	16	7	16	21
14	48	50	46	5	16	17
15	13	29	13		2	4
16	66	99	91	1	8	7
17	5	7	8			
18	24	57	50			
No. tested	246	307	278	247	309	277
Median Score	14.69	16.10	16.11	10.17	10.59	11.33

DISTRIBUTIONS

DIRECTIONS–VI				COMPOSITION			
	Grade				Grade		
Score	VI	VII	VIII	Score	VI	VII	VIII
0				0			
1				5			
2				10			
3				15	2	2	
4				20		2	
5				25	24	5	
6				30	60	35	11
7				35	77	65	22
8				40	35	84	47
9				45	25	54	72
10				50	22	41	63
11				55	8	22	47
12				60	1	3	13
13				65		1	7
14	1			70		1	1
15	2	1		75			
16	1	8	2	80			
17	11	13	6	85			
18	49	36	34	90			
19	138	175	177	95			
20	45	76	58		—	—	—
	—	—	—	No. tested	254	315	283
No. tested	247	309	277				
Median Score	19.43	19.55	19.55.	Median Score	37.66	42.89	49.27

Distributions

Visual Vocabulary–VIII

Score	Grade VI	VII	VIII
0			1
5		1	
10	1		
15			
20			1
25	1	2	1
30	5	1	
35	2	2	
40	2	1	2
45	8	2	
50	14	7	2
55	19	19	1
60	35	23	7
65	36	17	10
70	42	43	16
75	37	58	40
80	21	62	61
85	23	35	85
90	2	23	31
95		11	18
100		1	
No. tested	248	308	276
Median Score	70.12	78.10	84.67

Visual Vocabulary—XIII

Score	Grade VI	VII	VIII
0	1		
10	9	4	1
20	21	8	1
30	18	8	2
40	12	17	4
50	21	17	12
60	28	17	11
70	26	28	13
80	23	22	15
90	22	31	17
100	19	32	33
110	22	34	31
120	19	36	45
130	4	31	47
140	3	19	31
150		3	12
160			1
170		1	
No. tested	248	308	276
Median Score	75.38	100.63	119.35

DISTRIBUTIONS

OPPOSITES COMPOSITE A2–A3–A4				CONTROLLED ASSOCIATION COMPOSITE B1–B2–C1–C2–C3			
	Grade				Grade		
Score	VI	VII	VIII	Score	VI	VII	VIII
0	3	2		0			
10	1			5			1
20	1	2		10			
30	2	3		15	1		
40	2	1	1	20			
50	5			25	2		
60	6	1	1	30			
70	8	3		35	1		
80	4	4	1	40	2	2	
90	6	5	2	45	2	2	
100	14	5	2	50	1	3	2
110	20	5	1	55	3	4	
120	16	17	4	60	4	7	
130	12	11	10	65	13	4	2
140	18	18	12	70	11	8	8
150	26	20	11	75	9	7	6
160	21	20	7	80	13	17	14
170	13	24	19	85	20	16	12
180	19	23	21	90	18	15	25
190	16	18	13	95	20	39	24
200	12	20	24	100	29	30	22
210	6	21	18	105	37	38	37
220	6	17	23	110	40	68	58
230	7	18	19	115	24	50	57
240	1	12	14	120		2	3
250	1	7	17		—	—	—
260	1	7	13	No. tested	250	312	271
270	1	2	14				
280		4	8	Median Score	100.86	107.63	107.64
290		4	7				
300	1	7	4				
310		3	2				
320			3				
330			2				
340			1				
	—	—	—				
No. tested	249	304	274				
Median Score	152.50	134.78	214.44				

Distributions

Mixed Relations Composite D1–D2

Score	Grade VI	Grade VII	Grade VIII
0		2	
4	2		
8	2	3	
12	2	7	4
16	13	12	13
20	20	36	11
24	30	47	27
28	42	40	44
32	30	24	26
36	22	26	28
40	19	20	14
44	16	18	12
48	14	7	14
52	11	14	12
56	10	12	9
60	3	6	10
64	7	12	15
68	3	14	17
72	3	8	14
76	1	4	6
80			
No. tested	250	312	276
Median Score	33.86	33.40	37.86

Directions Composite X–VI–VII

Score	Grade VI	Grade VII	Grade VIII
0			
2			
4			
6			
8			
10			
12			
14			
16			
18			
20			
22			
24		2	
26	3		
28	1		
30		3	
32	1	3	4
34	7	3	3
36	9	14	11
38	32	13	14
40	36	26	28
42	53	53	34
44	53	73	63
46	37	66	67
48	12	29	32
50	3	20	18
52		3	3
54			
56			
No. tested	247	308	277
Median Score	43.30	45.01	45.41

DISTRIBUTIONS

VISUAL VOCABULARY COMPOSITE VIII–XIII				TRABUE COMPOSITE B–C–D–E			
	Grade				Grade		
Score	VI	VII	VIII	Score	VI	VII	VIII
				0			
0				5			
10				10			
20		1		15			
30	1			20			
40			1	25	1		
50	3			30	1		
60	7	3		35	8	8	
70	10	4		40	15	23	6
80	10	4	1	45	61	47	31
90	12	11	1	50	80	77	51
100	14	7	4	55	59	80	70
110	17	15	6	60	13	46	65
120	22	9	9	65	4	23	33
130	13	18	6	70		1	12
140	27	19	6	75			4
150	27	19	16	80			
160	17	23	7		—	—	—
170	14	24	21	No. tested	242	305	272
180	15	28	32				
190	17	34	25	Median Score	52.19	54.84	58.43
200	13	24	31				
210	5	27	35				
220	3	21	38				
230	1	11	22				
240		4	9				
250		2	6				
260							
	—	—	—				
No. tested	248	308	276				
Median Score	145.56	178.75	200.97				

DISTRIBUTIONS

ARITHMETIC COMPOSITE—Addition, Subtraction, Multiplication, Division, Problems I, Problems II

Score	Grade VI	VII	VIII
70		1	1
75	3	2	
80	4	3	
85	9	4	1
90	11	6	5
95	14	7	3
100	18	8	6
105	28	8	6
110	23	23	8
115	24	34	8
120	31	21	23
125	25	29	23
130	17	28	20
135	13	26	26
140	5	25	33
145	6	18	30
150	3	29	27
155	1	13	20
160	2	5	13
165	1	10	13
170		6	6
175		2	3
180			
185			
190			
195			
No. tested	238	308	275
Median	116.88	131.43	141.14

DISTRIBUTIONS

Score	ENGLISH Grade VI	COMPOSITE Grade VII	Grade VIII	Including all of the tests or groups of tests in the Total Composite except Arithmetic
190	1			
200	1			
210		2		
220	1			
230	2			
240	6	1		
250	4	4		
260	6	1		
270	8	4	2	
280	9	3		
290	8	6	2	
300	20	10	1	
310	16	7	4	
320	12	18	2	
330	23	15	12	
340	22	28	9	
350	24	21	8	
360	18	27	17	
370	19	26	15	
380	21	22	28	
390	11	23	17	
400	5	21	23	
410	8	20	34	
420	5	13	17	
430	3	13	17	
440	1	17	19	
450	1	10	20	
460		3	14	
470		2	11	
480		4	2	
490		1	3	
500		1	3	
510			2	
No. tested	255	323	282	
Median Score	344.77	375.58	410.29	

DISTRIBUTIONS

TOTAL COMPOSITE

Score	Grade VI	VII	VIII	Including:
				Arithmetic Composite
360		1		Trabue Composite
380	2	1		Alpha 2
400	4			Spelling
420	5	1		Composition
440	6	4	1	Opposites Composite
460	7	2	1	Controlled Association (B–C) Composite
480	19	11	1	
500	16	10	4	Mixed Relations Composite
520	26	6	6	Directions Composite
540	23	18	3	Visual Vocabulary Composite
560	17	17	7	
580	32	32	9	
600	29	27	16	
620	20	35	22	
640	19	33	29	
660	13	25	27	
680	9	23	31	
700	3	22	29	
720	2	18	22	
740	2	24	31	
760		5	16	
780		4	16	
800		2	9	
820		3	2	
No. tested	254	324	282	
Median Score	581.35	638.28	689.68	

PER CENTS OF OVERLAPPING FOR TESTS NOT INCLUDED IN TABLE II

Tests	VI on VIII	VI on VII	VII on VIII	VIII on VI	VIII on VII	VII on VI
A1	16.53	25.19	35.84	20.70	34.88	32.59
B1	42.98	37.40	56.42	40.11	55.89	36.78
B2	...	...	...	...	...	...
C1	27.19	40.07	37.66	29.48	37.70	41.01
C2	23.01	36.81	40.13	27.96	38.11	36.13
X	35.67	35.93	49.55	36.60	49.68	32.41
VI	43.36	43.36	50.08	42.64	50.31	43.12
PIII	7.87	21.63	26.84	...	28.86	...